Title

Discovery and verification of extracellular microRNA biomarkers for diagnostic and prognostic assessment of preeclampsia at triage

Journal

Science Advances, 9(51)

ISSN

2375-2548

Authors

Morey, Robert

Poling, Lara

Srinivasan, Srimeenakshi

et al.

Publication Date

2023-12-22

DOI

10.1126/sciadv.adg7545

Peer reviewed

HEALTH AND MEDICINE

Discovery and verification of extracellular microRNA biomarkers for diagnostic and prognostic assessment of preeclampsia at triage

Robert Morey[1]†, Lara Poling[1]†, Srimeenakshi Srinivasan[1], Carolina Martinez-King[1], Adanna Anyikam[1], Kathy Zhang-Rutledge[1]‡, Cuong To[1], Abbas Hakim[1], Marina Mochizuki[1], Kajal Verma[1], Antoinette Mason[1], Vy Tran[1], Morgan Meads[2], Leah Lamale-Smith[1], Hilary Roeder[1], Mariko Horii[2], Gladys A. Ramos[1], Peter DeHoff[1], Mana M. Parast[2] , Priyadarshini Pantham[1], Louise C. Laurent[1]*

We report on the identification of extracellular miRNA (ex-miRNA) biomarkers for early diagnosis and prognosis of preeclampsia (PE). Small RNA sequencing of maternal serum prospectively collected from participants undergoing evaluation for suspected PE revealed distinct patterns of ex-miRNA expression among different categories of hypertensive disorders in pregnancy. Applying an iterative machine learning method identified three bivariate miRNA biomarkers (*miR-522-3p/miR-4732-5p*, *miR-516a-5p/miR-144-3p*, and *miR-27b-3p/let-7b-5p*) that, when applied serially, distinguished between PE cases of different severity and differentiated cases from controls with a sensitivity of 93%, specificity of 79%, positive predictive value (PPV) of 55%, and negative predictive value (NPV) of 89%. In a small independent validation cohort, these ex-miRNA biomarkers had a sensitivity of 91% and specificity of 57%. Combining these ex-miRNA biomarkers with the established sFlt1:PlGF protein biomarker ratio performed better than either set of biomarkers alone (sensitivity of 89.4%, specificity of 91.3%, PPV of 95.5%, and NPV of 80.8%).

INTRODUCTION

Preeclampsia (PE) is a form of placental dysfunction that affects approximately 5 to 8% of pregnancies worldwide (*1*, *2*). PE manifests as hypertension and proteinuria, and in severe cases can lead to end-organ injury (*1*, *2*). Currently, the only intervention that can halt the progression of disease is delivery, and thus, many pregnancies affected by PE are delivered early to protect the health of the mother and fetus. As a result, PE is the leading cause of iatrogenic preterm birth.

Although it is thought that PE arises from abnormal trophoblast differentiation and invasion in early placental development (*3*–*5*), clinical manifestations do not arise until the second half of pregnancy. The most common symptom of PE is headache, followed by visual changes and abdominal pain. Typically, the diagnosis of PE is made when a pregnant patient displays new-onset hypertension (blood pressure >140/90 mmHg) and proteinuria (>0.30 g/24 hours) (*6*), but atypical cases where one of these features is absent or uninformative are not uncommon, particularly for women with preexisting conditions, such as chronic hypertension or kidney disease. A subset of patients with PE develop additional clinical features, including liver or kidney dysfunction, low platelet counts, cerebral edema, seizures, or placental abruption. Given the broad array of clinical presentations for PE, it can be difficult to differentiate the early signs and symptoms of PE from other conditions.

Currently, early diagnosis of and/or risk assessment for the later development of PE is problematic due to the lack of assays that are highly specific for this disease. Accurate evaluations are important when planning the intensity of pregnancy surveillance or determining the timing of delivery. If delivery is induced too early, the neonate may be unnecessarily exposed to complications associated with prematurity. However, if the decision to deliver is made too late, the mother and neonate may be exposed to an increased risk of severe manifestations of PE, which can lead to serious morbidity or death. This dilemma frequently leads to expensive and lengthy hospital stays for patients at high risk of PE. The management of PE could be substantially improved, and unnecessary hospitalizations could be avoided, if an assay were developed to diagnose early PE rapidly and accurately and/or assess the risk of severe PE.

At present, a highly sensitive and specific assay for early diagnosis or prognostic evaluation of PE does not exist (*7*), although multiple commercial protein biomarker tests for PE have been reported. The most accurate of these tests are designed to detect two proteins in the blood: placental growth factor (PlGF) and soluble FMS-like tyrosine kinase 1 (sFlt1). PlGF is important in the development of a healthy placenta, as it promotes the formation of new blood vessels, while sFlt1 inhibits the function of PlGF (*8*). In women with PE, PlGF levels may be decreased, while sFlt1 levels may be increased (*8*). The level of these two proteins is reported as a ratio of sFlt1:PlGF. As the ratio of sFlt1:PlGF increases, so does the risk of preterm PE.

The primary limitation of the sFlt1:PlGF ratio is its relatively low positive predictive value (PPV), which numerous studies have reported to be between 30 and 65% (*9*–*12*). Still, it is currently the most promising protein biomarker screening tool for PE. The benefit of the sFlt1:PlGF ratio is its negative predictive value

[1]Department of Obstetrics, Gynecology and Reproductive Medicine, University of California, San Diego, La Jolla, CA, USA. [2]Department of Pathology, University of California, San Diego, La Jolla, CA, USA.
*Corresponding author. Email: llaurent@ucsd.edu
†These authors contributed equally to this work.
‡Present address: Obstetrix Maternal-Fetal Medicine Specialists of Houston, Houston, TX, USA.

(NPV), which is helpful in ruling out PE and identifying women at low risk for development of PE among those with signs or symptoms of PE. While sFlt1:PlGF may be helpful in certain scenarios, the diagnosis of PE still depends on standard clinical signs and symptoms. Studies have yet to demonstrate that pregnancy outcomes are improved by screening suspected cases of PE using the sFlt1:PlGF ratio (*8*).

Although the sFlt1:PlGF ratio has been available for almost 20 years, it was only recently approved by the U.S. Food and Drug Administration (FDA) and at this time is excluded from the PE screening practices recommended by the U.S. Preventative Services Task Force, the National Institute for Health and Care Excellence, the American College of Obstetricians and Gynecologists, and the Society of Obstetricians and Gynecologists of Canada. Collectively, these organizations find limited clinical utility in the commercially available PE biomarker tests and have requested the development of better screening biomarkers and tools for risk prediction (*7*).

To improve upon existing assays, many research efforts have evaluated proteins other than sFlt1 and PlGF (*13*–*21*). Some studies have moved beyond protein biomarkers altogether to nucleic acid–based laboratory tests. These studies initially focused on long noncoding RNA molecules, cell-free DNA, and cellular mRNA (*22*–*25*), but over the past decade, an increasing number of studies have explored the utility of extracellular microRNAs (ex-miRNAs) as diagnostic or prognostic biomarkers for PE (*26*–*43*).

Many of these biomarker studies have used a candidate approach, measuring levels of specific biomolecules based on results of previous studies demonstrating that they are expressed specifically by the placenta and/or play a role in placental development or function. While this approach has yielded notable results in several published studies, none of these studies have led to the development of a clinical assay. Thus, both the utility of nucleic acid biomarkers for PE and investigation of their roles in the pathophysiology of this disease are still in early stages.

Since it has not been established that only placental biomolecules are perturbed in women at elevated risk of developing PE, we reasoned that a comprehensive unbiased approach might reveal biomarkers for PE. Therefore, the primary aim of this study was to use an unbiased transcriptomic approach to discover ex-miRNA biomarkers for the early diagnosis of PE in a cohort of pregnant women being evaluated for signs and symptoms of PE. Post hoc analysis was also performed to explore the utility of ex-miRNA biomarkers for prediction of severity of PE in those who developed this complication.

RESULTS

Characteristics of the study population

A total of 131 subjects were recruited and enrolled. Of these, one subject was lost to follow-up and seven were excluded during quality control of the small RNA sequencing (RNA-seq) data due

Table 1. Study subject demographics and clinical characteristics. Demographics and clinical characteristics of subjects used for statistical analysis. *P* values represent statistical significance for characteristics between cases and controls within the discovery and verification cohorts, and between the discovery and verification cohorts (rightmost column). GA, gestational age; GABD, gestational age at blood draw; NICU, neonatal intensive care unit; small for GA, birthweight <10th percentile for GA.

	Discovery (*n* = 82)			Verification (*n* = 41)			
	PE Dx Case = 48	All Other Dx Control = 34	*P*	PE Dx Case = 23	All Other Dx Control = 18	*P*	*P*
Mean maternal age (years)	31.9 ± 6.5	31.2 ± 5.4	0.6	29 ± 6	33.2 ± 5.2	0.03	0.67
Median gravidity	2 ± 2.3	2 ± 2	0.3	2 ± 2.5	2 ± 1.3	0.39	0.75
Median parity	0.5 ± 1.2	1 ± 1.6	0.06	0 ± 1.4	1 ± 1	0.26	0.7
Mean first BMI	29.9 ± 7.2	28.9 ± 6.9	0.6	32.8 ± 8	27.6 ± 8.7	0.1	0.78
Mean last BMI	33.0 ± 7.3	32.7 ± 6.5	0.81	36.8 ± 8.4	31.6 ± 9	0.08	0.27
No. of subjects with diabetes	18 (37.5%)	9 (26.5%)	0.35	8 (34.8%)	8 (44.4%)	0.75	0.55
Race/ethnicity							
Hispanic	21 (43.8%)	13 (38.2%)		11 (47.8%)	6 (33.3%)		
White non-Hispanic	18 (37.5)	17 (50%)		5 (21.7%)	7 (38.9%)		
Asian	6 (12.5%)	0	<0.01	1 (4.3%)	4 (22.2%)	<0.01	<0.01
African American	3 (6.3%)	0		2 (8.7%)	0		
Other	0	4 (11.8%)		4 (17.4%)	1 (5.6%)		
Median GABD (days)	238 ± 30.2	245 ± 24.0	0.25	244 ± 25.7	241 ± 25.6	0.79	0.66
Median GA delivery (days)	247.5 ± 26.9	273 ± 12.4	<0.01	249 ± 25.8	272 ± 8.1	<0.01	0.49
No. of preterm deliveries	12 (25%)	4 (11.8%)	0.16	5 (21.7%)	5 (27.8%)	0.72	0.64
Mean birthweight (g)	2188 ± 872	3240 ± 516.5	<0.01	2345 ± 1005	3258.9 ± 396.3	0.001	0.48
No. of NICU	37 (80.4%)	9 (27.3%)	<0.01	16 (69.6%)	2 (11.1%)	<0.01	0.18
No. of small for GA	24 (52.2%)	3 (88%)	<0.01	10 (43.5%)	1 (5.6%)	0.01	0.54

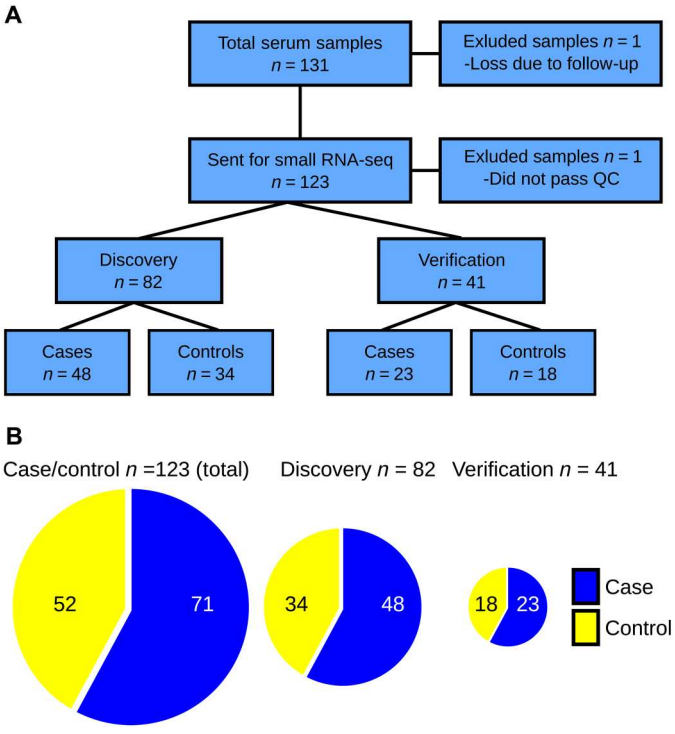


Fig. 1. Study population characteristics. (**A**) Flow chart depicting the exclusion of seven subjects due to insufficient RNA quality and one subject due to loss at follow-up. The remaining 123 samples were submitted for analysis (71 cases, 52 controls). (**B**) Pie charts depicting the breakdown of cases and controls in the cohort.

to insufficient miRNA complexity or low read depth (Fig. 1A). Cases were defined as pregnancies with an adjudicated PE diagnosis (n = 71), and controls (n = 52) were from participants with either hypertension that did not meet the criteria for PE or normal (non-hypertensive) pregnancy outcomes. The pass-filter 71 cases and 52 controls were subjected to further analysis. Cases and controls were divided into discovery (n = 82; 48 PE cases and 34 controls) and verification (n = 41; 23 PE cases and 18 controls) cohorts at a ratio of 2:1. The discovery and verification cohorts were matched by fraction of cases and controls (Fig. 1B) and gestational age at blood draw (GABD), which ranged from 22.6 to 39.1 weeks for cases and 32.6 to 40.4 weeks for controls. Data file S1 lists all subjects and includes assignment to discovery and verification cohorts, assignment as cases and controls, and detailed adjudicated diagnoses. Data file S8 contains sample-level metadata.

Comparisons for demographic and clinical factors between the discovery and verification cohorts, and between the cases (PE Dx) and controls (All Other Dx) were then performed (Table 1). Significant differences ($P \leq 0.05$) were observed for median gestational age (GA) at delivery, mean birthweight, small for gestational age (SGA; birth weight of <10%), and admission to the neonatal intensive care unit (NICU) between the cases and controls in both the discovery and verification cohorts; these differences are expected given the known associations between PE and outcomes such as iatrogenic delivery and SGA neonates. There was also a significant difference in race/ethnicity between cases and controls, with higher percentages of Hispanic and non-white subjects in the case compared to the control category. All other factors were similar between cases and controls for both the discovery and verification cohorts. The only significant difference between the discovery and verification cohorts was seen for race/ethnicity, with the predominant difference being a larger proportion of white non-Hispanic participants in the discovery cohort.

Generation and pre-processing of ex-miRNA data

As detailed in Materials and Methods, exRNA was extracted from maternal serum samples and then subjected to small RNA-seq. The small RNA-seq data were trimmed, cleaned, and mapped using the exceRpt pipeline (*44*), and the raw count data were used for quality control, removing samples with total miRNA counts <500,000 or complexity <300 miRNAs with at least 10 raw counts

Table 2. Top 10 ranked univariate extracellular miRNA (ex-miRNA). Ex-miRNAs were ranked according to chi-square *P* value in the discovery cohort, and the top 100 were then evaluated in the verification cohort. The top 10 univariate ex-miRNAs ranked by chi-square *P* value in the verification cohort that also had a discovery chi-square rank <100 are shown. Discovery and verification rank and chi-square *P* value are shown along with the verification AUC at the lower 25th and upper 75th percentiles and whether the expression of the ex-miRNA is up- or down-regulated in the non-preeclampsia (PE) versus PE samples.

miRNA	Training rank (chi-square *P*)	Discovery (chi-square *P*)	Discovery rank	Discovery (chi-square *P*)	Discovery AUC lower 25	Discovery AUC upper 75	Non-PE vs. PE
hsa.miR.4732.5p	64	0.299	4	0.039	0.008	0.2	Up
hsa.miR.363.3p	22	0.188	3	0.039	0.007	0.201	Up
hsa.let.7b.5p	1	0.015	5	0.047	0.012	0.22	Up
hsa.miR.423.5p	2	0.041	2	0.036	0.011	0.221	Up
hsa.miR.144.3p	75	0.317	11	0.077	0.046	0.307	Up
hsa.miR.1323	82	0.326	9	0.075	0.683	0.945	Down
hsa.miR.361.5p	68	0.307	8	0.07	0.708	0.958	Down
hsa.miR.512.3p	31	0.206	10	0.076	0.725	0.961	Down
hsa.miR.518e (519a,519b,519c,522,523).5p	96	0.352	7	0.063	0.745	0.975	Down
hsa.miR.516a.5p	30	0.205	1	0.034	0.801	0.991	Down

in at least 50% of the samples. Transformation and batch normalization was conducted using the PEER package as described in Materials and Methods.

Identification of univariate ex-miRNA biomarkers for early diagnosis of PE

The PEER package was used to identify candidate univariate biomarkers composed of single ex-miRNAs (*45*). The ex-miRNAs were ranked according to chi-square *P* value in the discovery cohort, and the top 100 were then evaluated in the verification cohort. The top 10 univariate ex-miRNAs ranked by chi-square *P* value in the verification cohort that also had a discovery chi-square rank <100 are shown in Table 2. For the verification cohort, in addition to the *P* value, we used PEER to compute the 25th and 75th percentile areas the curve (AUCs) for these candidate ex-miRNA biomarkers, which was done by subsampling the samples 1000 times. We discovered that although only a subset of the candidate ex-miRNAs reached statistical significance according to the chi-square *P* value, all the candidate ex-miRNAs that were more highly expressed in non-PE samples compared to PE samples showed a 75th percentile AUC of <0.35, and all the candidates that were more highly expressed in PE samples compared to non-PE samples showed a 25th percentile AUC of >0.65.

The expression of these top 10 candidate univariate ex-miRNA biomarkers was then examined on a per-sample basis (Fig. 2 and data file S2). As can be seen in Fig. 2, which shows all patients, there is a strong separation of the samples by overall diagnosis (case versus control), by detailed diagnoses (Severity Dx), and by interval between blood draw and PE diagnosis, but not by cohort (discovery versus verification) or GABD. Superimposed PE cases appear to be scattered across several clusters of samples, some of which are composed mostly of other PE samples and others of hypertensive non-PE samples; this may be because it is clinically challenging to diagnose superimposed PE and/or that this condition contains molecular features of both chronic hypertension and PE. We note that 4 of these 10 candidate miRNAs are encoded on Chr19 and are more highly expressed in PE compared to non-PE, and 3 are on Chr17 and are more highly expressed in non-PE.

Identification of bivariate ex-miRNA biomarkers for early diagnosis of PE

The PEER package was used to identify candidate bivariate biomarkers composed of ratios of the log values for pairs of ex-miRNAs. To do this, the bivariate features were ranked in the discovery cohort according to a composite score derived from four metrics: correlation coefficient, chi-square *P* value, 25th percentile AUC, and 75th percentile AUC (see Materials and Methods for details). The top 1000 bivariate features from the discovery cohort were then evaluated in the verification cohort, and the features with a 25th percentile AUC of at least 0.7 were selected, yielding 110 candidate bivariate biomarkers that passed verification (data file S3).

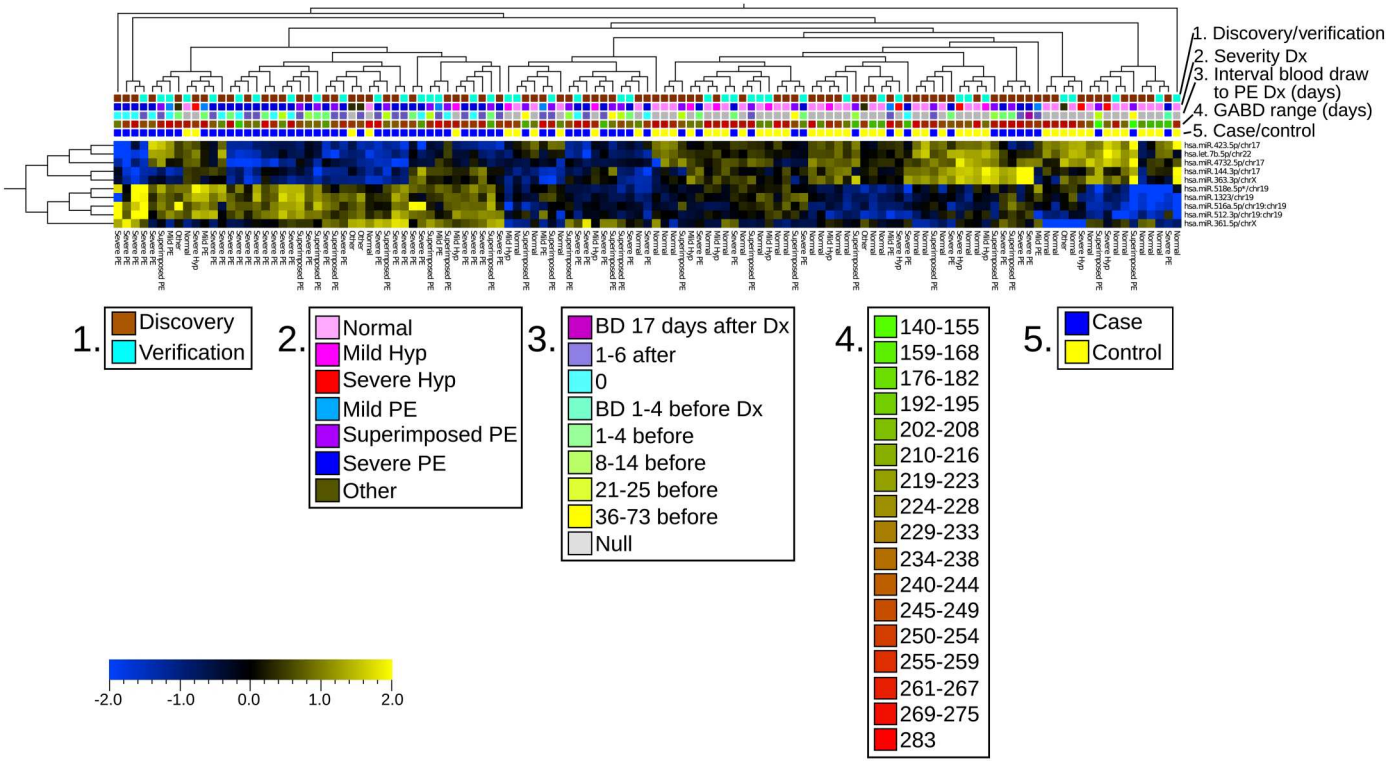


Fig. 2. Expression of the top 10 candidate univariate extracellular miRNA (ex-miRNA) biomarkers. Heatmap displaying the expression of the top 10 candidate univariate ex-miRNA biomarkers. Both ex-miRNAs (*y* axis) and samples (*x* axis) are clustered using hierarchical clustering. Color bars at the top of the heatmap display sample metadata and show strong separation of the samples by overall diagnosis (case versus control; $n = 71$ and $n = 52$, respectively), by detailed diagnoses, and by interval between blood draw and preeclampsia (PE) diagnosis, but not by cohort (discovery versus verification; $n = 82$ and $n = 41$, respectively) or gestational age at blood draw (GABD). No patients were included in multiple classifications.

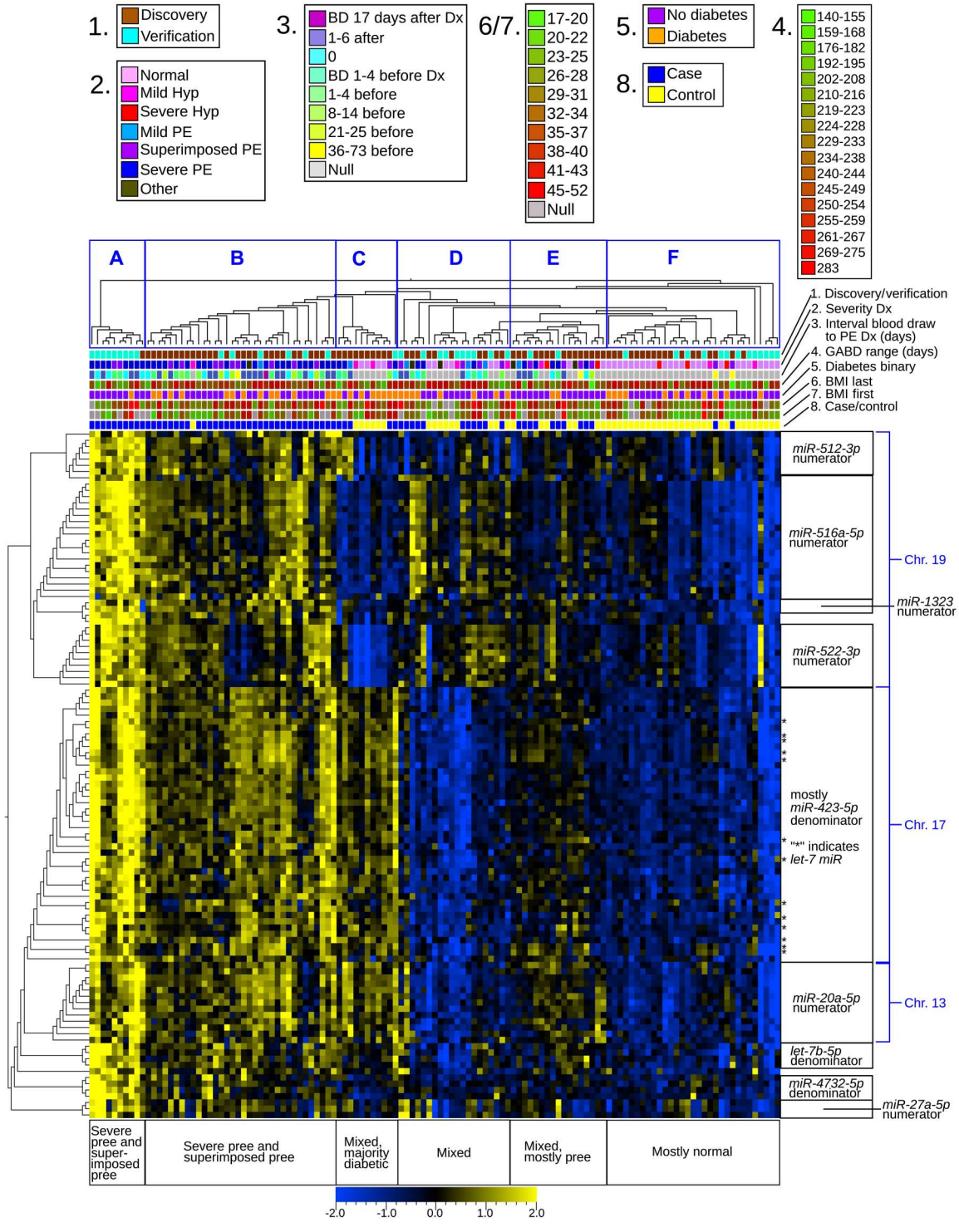


Fig. 3. Expression of the 110 verified candidate bivariate extracellular miRNA (ex-miRNA) biomarkers. Heatmap displaying the log ratio of the 110 verified candidate bivariate ex-miRNA biomarkers. Both bivariate ex-miRNAs (*y* axis) and samples (*x* axis) are clustered using hierarchical clustering. Color bars at the top of the heatmap display sample metadata and show strong separation of the samples by overall diagnosis (case versus control; $n = 71$ and $n = 52$, respectively), by detailed diagnoses, and by interval between blood draw and preeclampsia (PE) diagnosis, but not by cohort (discovery versus verification; $n = 82$ and $n = 41$, respectively) or gestational age at blood draw (GABD). No patients were included in multiple classifications. Blue boxes with letters mark sample clusters discussed in the text. Bivariate ex-miRNA biomarkers clustered by the presence of a particular ex-miRNA biomarker in the ratio and by location in the genome as shown on the *y* axis.

The expression of these top 110 candidate bivariate ex-miRNA biomarkers was then examined on a per-sample basis (Fig. 3 and data file S4). As can be seen in Fig. 3, which shows all patients, there is strong separation of the samples by overall diagnosis (case versus control), by detailed diagnoses, and interval between blood draw and PE diagnosis, but not by cohort (discovery versus verification), GABD, maternal diabetes, or body mass index (BMI).

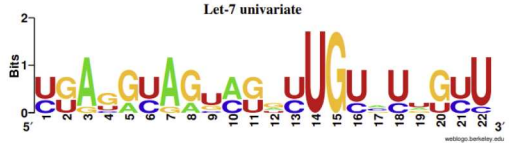
Inspecting the miRNA composition of these top 110 candidate bivariate biomarkers, we observe that several of the candidate univariate biomarkers are present, with the Chr19 biomarkers significantly enriched in the numerators [$P < 0.01 \times 10^{28}$, cumulative distribution function (CDF) of the hypergeometric distribution] and the Chr17 biomarkers significantly enriched ($P < 0.01 \times 10^{36}$, CDF of the hypergeometric distribution) in the denominators. Certain individual miRNAs appear in multiple bivariate biomarkers, and these sets of bivariate biomarkers tend to cluster together by hierarchical clustering (Fig. 3). When visualized by chromosomal location, it is apparent not only that certain individual miRNAs are enriched in the set of 110 candidate bivariate biomarkers but also that certain genomic locations are overrepresented, particularly on chromosomes 17 and 19 (fig. S1).

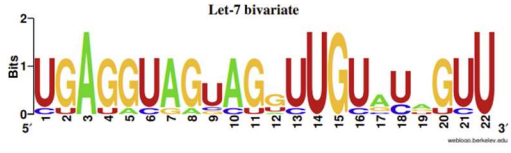
Compared to the candidate univariate biomarkers, the candidate bivariate biomarkers appear to better separate subjects into phenotypic subgroups, as indicated by the text in the boxes below the heatmap in Fig. 3. Going from left to right, cluster A consists mostly of samples from subjects who developed severe PE, along

A

PE regulated	**Let-7: Univariate Top 100**	
Up-regulated	*let-7g-5p*	UGAGGUAGUAGUUUGUACAGUU
Up-regulated	*let-7i-5p*	UGAGGUAGUAGUUUGUGCUGUU
Up-regulated	*let-7d-3p*	CUAUACGACCUGCUGCCUUUCU
Up-regulated	*let-7f.2-3p*	CUAUACAGUCUACUGUCUUUCC
Up-regulated	*let-7i-3p*	CUGCGCAAGCUACUGCCUUGCU
Down-regulated	*let-7b-5p*	UGAGGUAGUAGGUUGUGUGGUU
Down-regulated	*let-7c-5p*	UGAGGUAGUAGGUUGUAUGGUU
Down-regulated	*let-7e-5p*	UGAGGUAGGAGGUUGUAUAGUU

	Let-7: Bivariate Top 110	
Numerator	*let-7d-3p*	CUAUACGACCUGCUGCCUUUCU
Numerator	*let-7a-5p*	UGAGGUAGUAGGUUGUAUAGUU
Numerator	*let-7f-5p*	UGAGGUAGUAGAUUGUAUAGUU
Numerator	*let-7g-5p*	UGAGGUAGUAGUUUGUACAGUU
Numerator	*let-7i-5p*	UGAGGUAGUAGUUUGUGCUGUU
Denominator	*let-7b-5p*	UGAGGUAGUAGGUUGUGUGGUU
Denominator	*let-7c-5p*	UGAGGUAGUAGGUUGUAUGGUU
Denominator	*let-7e-5p*	UGAGGUAGGAGGUUGUAUAGUU





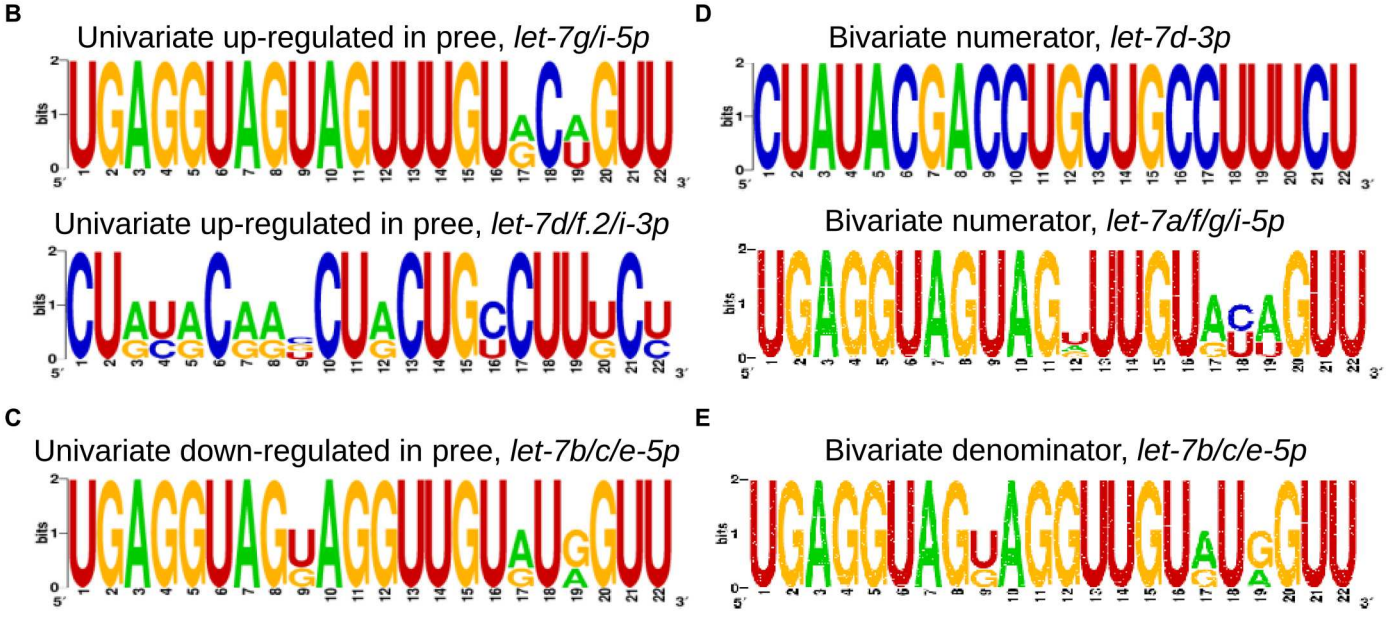


Fig. 4. Let-7 family clusters and sequence logos. (**A**) Table showing the let-7 family extracellular miRNA (ex-miRNA) found in the top 100 univariate biomarkers (top) and the top 110 bivariate biomarkers (bottom). The univariate table is divided by let-7 ex-miRNAs that are either up-regulated or down-regulated in the preeclampsia (PE) samples and by 5p and 3p. The bivariate table is divided by the let-7 ex-miRNAs found in either the numerator or the denominator of the bivariate ratio. The sequence logos to the right of the table show a graphical representation of the sequence conservation of nucleotides in the ex-miRNAs. (**B**) Sequence logos showing a graphical representation of the two let-7 univariate 5p sequences (top) and the three *let-7-3p* sequences (bottom) up-regulated in our PE samples. (**C**) Sequence logo showing a graphical representation of the three let-7 univariate 5p sequences down-regulated in our PE samples. (**D**) Sequence logos showing a graphical representation of the one *let-7-3p* (top) and four *let-7-5p* (bottom) sequences found in the numerator of the bivariate ratios. (**E**) Sequence logo showing a graphical representation of the three *let-7-5p* sequences found in the denominator of the bivariate ratios.

with two cases of superimposed PE, which have uniformly high expression of all the candidate bivariate biomarkers. The interval between GABD and initial diagnosis of any form of PE in this group ranges from 14 days before to 6 days after, suggesting that this pattern of biomarker expression indicates a current or imminent clinical diagnosis of PE with severe features. Cluster B contains a relatively even mixture of cases of severe PE and superimposed PE, along with one case of severe hypertension and one case of mild PE. The samples in cluster B express moderate levels of all the candidate bivariate biomarkers. The interval between GABD and initial diagnosis of any form of PE in cluster B is generally larger than in cluster A and ranges from 73 days before to 6 days after, suggesting that the pattern of biomarker expression in cluster B may serve as an early predictor of PE with severe features. Cluster C shows uniform and moderate expression of the bivariate biomarkers in the lower portion of the heatmap (Fig. 3, *miR-20a-5p* numerator, *miR-423-5p* denominator, let-7 family) and mostly low expression of the biomarkers in the upper portion of the heatmap (Fig. 3, *miR-512-3p* numerator, *miR-516a-5p* numerator, *miR-522-3p* numerator) and consists of a mixture of severe, mild, and superimposed PE, mild hypertension, and normal (nonhypertensive) cases. However, it does appear that all but one of the PE cases shows moderately high expression of the biomarkers with *miR-512-3p* in the numerator, suggesting that cluster C is composed of two subclusters, one for PE and one for non-PE. Clusters D and E contain samples with a broad mix of diagnoses, and a small number of cases of severe PE, but mostly mild and superimposed PE, mild and severe hypertension, and normal (nonhypertensive). Cluster D samples show low-to-moderate expression of the biomarkers in the upper portion of the heatmap, and essentially no expression of the biomarkers in the lower portion of the heatmap, while cluster E shows low expression of the biomarkers in the lower portion of the heatmap and very low expression of the biomarkers in the upper portion of the heatmap. Cluster F shows little to no expression of any of the bivariate biomarkers and is composed of mostly normal samples, with a few mild and severe hypertension and very few superimposed PE samples. Given that there was a cluster of bivariate biomarkers that was enriched for let-7 family miRNAs in both the numerator and denominator, we examined the expression of the let-7 family miRNAs in our dataset in more detail (Fig. 4A). We observed that in the top 110 bivariate biomarkers, *let-7b-5p*, *let-7c-5p*, and *let-7e-5p* occurred in the denominator, which is consistent with our univariate findings that these miRNAs are expressed at lower levels in PE compared to non-PE samples. Although the 3p/5p arms are not perfectly concordant between the bivariate and univariate datasets, the *let-7a/d/f/g/i* miRNAs are consistently found in the numerators and more highly expressed in PE than in non-PE. When we examine the sequences of these miRNAs, we find that for the 5p sequences, a C in position 18 is associated with higher expression in PE as a univariate biomarker (Fig. 4B)/bivariate numerator (Fig. 4D), while a U in position 18 is associated with lower expression in PE as a univariate biomarker (Fig. 4C)/bivariate denominator (Fig. 4E). The 3p sequences were only found to have higher expression in PE as univariate biomarkers (Fig. 4B)/bivariate numerator (Fig. 4D); we note that the sequence concordance among these *let-7-3p* family members was markedly lower than for the 5p family members.

Expression patterns of candidate biomarkers across detailed diagnoses

Given that we had detailed adjudicated diagnoses available [severe PE: PE with severe features, including HELLP (hemolysis, elevated liver enzymes, and low platelets); superimposed PE; mild PE: PE without severe features; other: atypical PE, mild PE with SGA, severe gestational hypertension with SGA, gestational proteinuria, gestational thrombocytopenia, severe chronic hypertension with headache and SGA; mild hypertension: mild chronic hypertension or mild gestational hypertension; severe hypertension: severe chronic hypertension or severe gestational hypertension; and normal/nonhypertensive], we examined the expression of our candidate univariate and bivariate biomarkers across these more granular categories.

Ordering the detailed diagnoses from more severe to less severe (i.e., severe PE on the left to normal/nonhypertensive on the right), we see, as might be expected that for most of the 10 candidate univariate biomarkers that there is a smooth progression from either higher to lower, or lower to higher, expression (fig. S2). However, interestingly, there were a few miRNAs for which the levels in samples from severe hypertension are more similar to normal/nonhypertensive and mild hypertension are more similar to PE (*let-7b-5p*, *miR-144-3p*, *miR-361-5p*, *miR-423-5p*).

For the 110 candidate bivariate biomarkers, we performed hierarchical, agglomerative clustering, using a weighted method with a Pearson correlation distance metric to identify 10 different patterns of expression across the detailed diagnoses. The patterns for three of the largest clusters are shown in Fig. 5 (A to C). Most miRNAs in these three clusters are found in just two regions on the genome, one on chromosome 17 and the other on chromosome 19. Clusters 2 and 3 (Fig. 5, A and B) were significantly enriched for miRNAs found on chromosome 19 in their numerator ($P < 0.005$ CDF of the hypergeometric distribution). Ninety percent of the miRNAs found in the numerator of cluster 2 and all the miRNAs in the numerator of cluster 3 come from the primate-specific microRNA cluster C19MC. The largest cluster, cluster 7 (Fig. 5C), was significantly enriched ($P < 0.00005$, CDF of the hypergeometric distribution) for the chromosome 17 miRNA *miR-423-5p*, found in the denominator of over 80% of bivariates in the cluster. This miRNA is in the same cluster as *miR-4732-5p*, which was enriched in cluster 2 (Fig. 5A). Compared to both the normal/nonhypertensive and the severe hypertension groups, this cluster contains high ratios of all the PE groups as well as the mild hypertension group.

Using a recently reported approach for estimation of the fractional contributions of cell/tissue sources to miRNAs in maternal serum *(45)*, we identified the most likely cell/tissue sources for our 10 candidate univariate and 110 candidate bivariate biomarkers. As shown in Fig. 6, the large majority of these candidates likely originate from the liver, placenta, platelets, or red blood cells (RBCs). Most of the miRNAs attributed to the placenta are in the "Up in PE"/Numerator categories, compared to the other three major cell/tissue sources, where the "Up in PE"/Numerator and "Down in PE"/Denominator categories are quite evenly balanced. The miRNAs that were up-regulated in PE were found only in the placenta and, to a lesser extent, the liver, whereas those that were deemed to be up-regulated in our control patients were found in platelets or RBCs. These analyses suggest that specific miRNA clusters influence pregnancy-related functions and that they are dysregulated in PE placentas. Moreover, the dysregulation of miRNAs

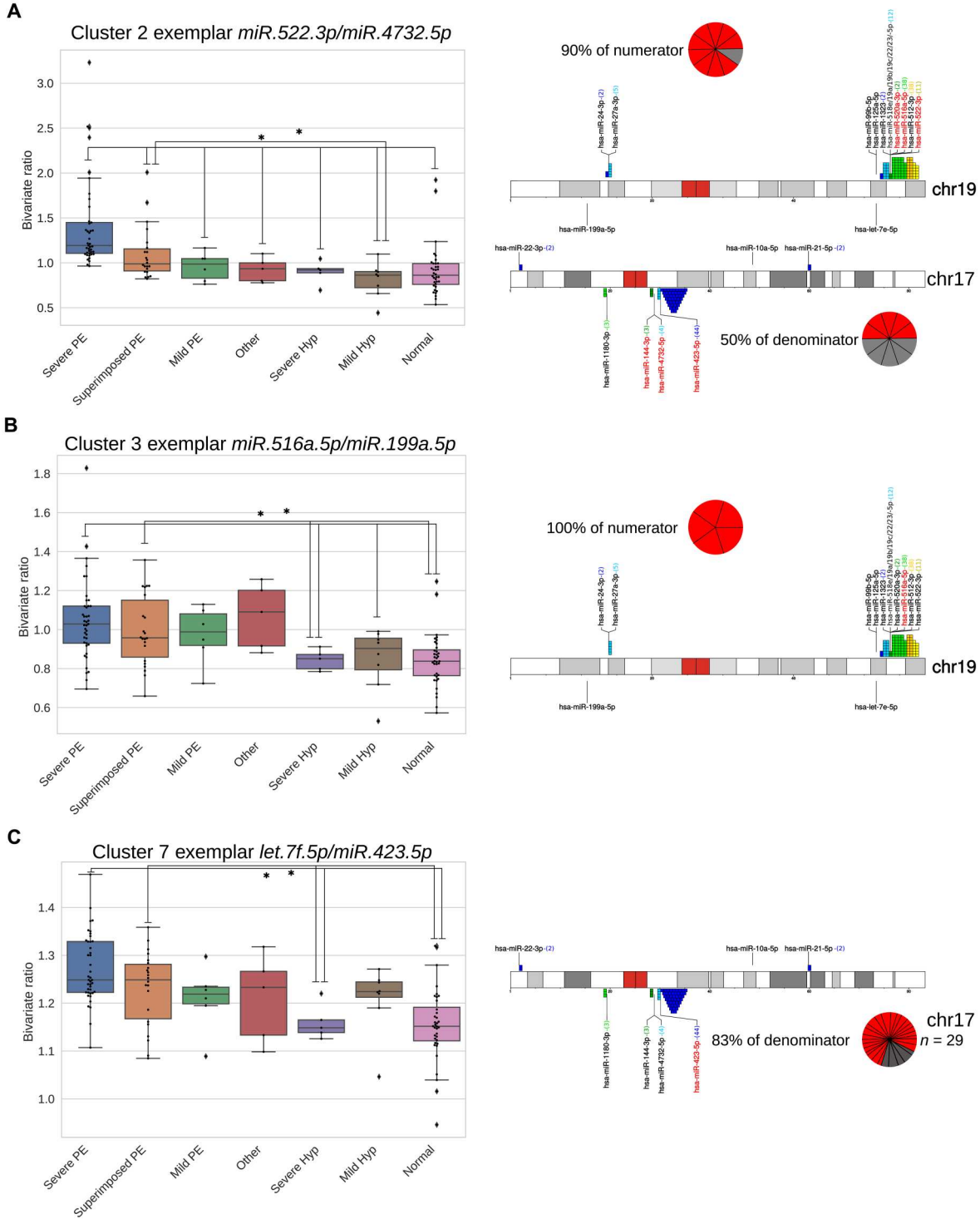


Fig. 5. Clustering of the 110 verified candidate extracellular miRNA (ex-miRNA) biomarkers. Three of the largest clusters (out of 10) formed using hierarchical, agglomerative clustering using a weighted method and Pearson correlation distance metric. Boxplots (left) show the ratios of an exemplar ex-miRNA. Samples are grouped along the *x* axis by their detailed diagnosis. Lines across boxes represent the median; boxes contain the 25th to 75th percentile. Chromosome ideogram with locations of miRNA clusters found in the 110 verified candidate biomarkers (right). Colored numbers in () at the end of the miRNAs represent the number of ex-miRNAs in the 110 verified ratios potentially present at that location. miRNAs in red font are those present in the cluster displayed in box plot. Pie graphs show the percentage of the cluster's numerator or denominator in the displayed chromosome. (**A**) Boxplot of cluster 2 exemplar *miR-522-3p/miR-4732-5p* (left) and location of a percentage (pie chart) of cluster 2's ex-miRNA bivariates (right). (**B**) Boxplot of cluster 3 exemplar *miR-516a-5p/miR-199a-5p* (left) and location of a percentage (pie chart) of cluster 3's ex-miRNA bivariates (right). (**C**) Boxplot of cluster 7 exemplar *let-7f-5p/miR-423-5p* (left) and location of a percentage (pie chart) of cluster 7's ex-miRNA bivariates (right). Statistical significance bars with "*" indicates statistically significant difference ($P < 0.01$) between groups using *t* test.

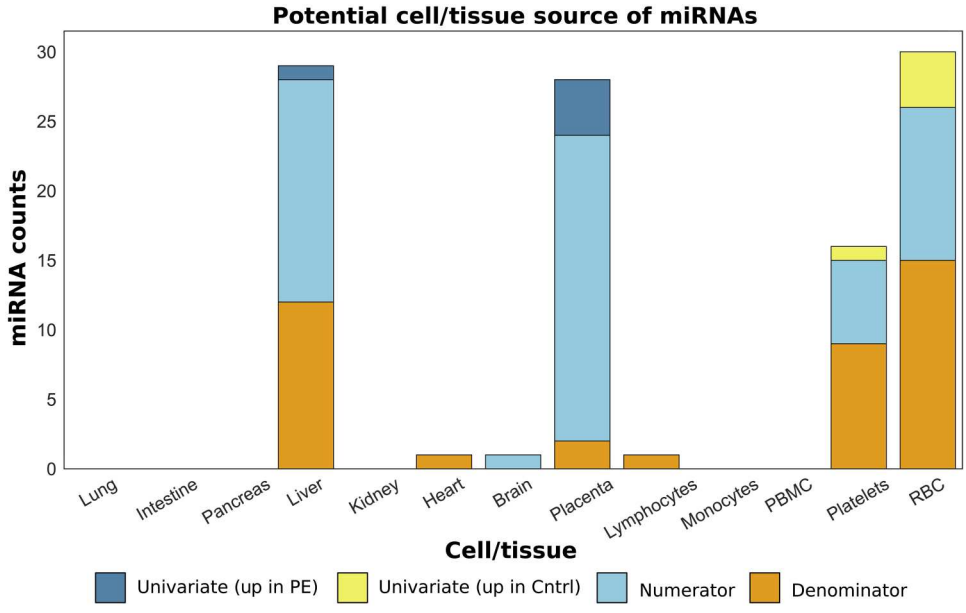


Fig. 6. Potential cell/tissue sources for our candidate extracellular miRNA (ex-miRNA) biomarkers. Stacked bar graph of the cell and tissue sources for our candidate ex-miRNA biomarkers. PE, preeclampsia; RBC, red blood cell.

can be traced to the placenta or to sources in contact with the maternal blood supply.

Identification of a small panel of candidate bivariate biomarkers for improved discrimination between cases and controls

To improve discrimination between cases and controls, we identified a small panel of candidate bivariate biomarkers from the set of 110 candidate bivariate biomarkers from the verification cohort using extreme gradient boosting. We performed an iterative approach where we selected the top bivariate, in terms of importance, when running the algorithm on the entire dataset and then removed all the samples that had a bivariate ratio that was greater than 90% of the controls. We then re-ran the algorithm an additional two times, correctly identifying all but five PE cases, three of which were diagnosed as superimposed PE and one mild PE (Fig. 7A). The bivariate with the highest importance in our first iteration was *miR-522-3p/miR-4732-5p* (Fig. 7A, top heatmap). This bivariate was interestingly also the exemplar bivariate for cluster 2 (Fig. 5A), which was a cluster of bivariates that clearly separated severe PE patients and normal patients. This single bivariate was able to separate over 70% of all severe PE diagnosed cases from controls. The bivariate with the highest importance score on the second iteration was *miR-516a-5p/miR-144-3p* (Fig. 7A, middle heatmap). This bivariate contained *miR-144-3p*, a chromosome 17 miRNA that comprised half of the denominators in cluster 1 (Fig. 5A), and *miR-516a-5p*, a chromosome 19 miRNA that comprised 100% of the numerators in cluster 3 (Fig. 5B), which was a cluster of bivariates that differentiated superimposed PE cases from controls. The third iteration indicated that *miR-27b-3p/let-7b-5p* had the highest importance score (Fig. 7A, lower heatmap). This bivariate was found to differentiate mild PE cases from controls to a larger degree ($P < 0.0015$, t test) than all but one of the other 110 verified bivariate biomarkers and was responsible for selecting over half of the mild PE cases from the remaining cases and controls. This iterative panel of three bivariate biomarkers achieved a PPV of 55% at a sensitivity of 93%, specificity of 79%, a positive likelihood ratio (+LR) of 4.43, and a negative likelihood ratio (−LR) of 0.09, given a 57.7% prevalence of PE in our study. We noted that the GA and BMI did not appear to be driving the selection of any of three bivariate biomarkers (Fig. 7A and data file S5). Additionally, we checked these three bivariate markers for correlation with patients′ urine protein-to-creatinine (PC) ratio, 24-hour urine protein levels, platelet count, uric acid, aspartate transaminase (AST), alanine transaminase (ALT), and systolic and diastolic blood pressure (highest values on the day closest to the date of blood draw) and found very little correlation (maximum correlation <0.39). We next performed principal components analysis using these three selected bivariate biomarkers (Fig. 7B) and a randomly chosen set of three bivariate biomarkers from the top ranked 1000 bivariate biomarkers (Fig. 7C). We found that, as expected, the three selected biomarkers clearly separated the cases and controls on the first three principal components, in the order that they were selected. The first three principal components of the random three biomarkers did not appear to differentiate between cases and controls. To verify that this iterative panel of three bivariate biomarkers would work on a separate cohort, we used the ratio cutoff values iteratively for each of the three bivariate biomarkers on a separate cohort of 11 cases and 7 controls processed separately through PEER. This independent validation cohort had a PPV of 85% at a sensitivity of 91% and specificity of 57% (fig. S3 and data file 6). These results suggest that a small panel of bivariate biomarkers can be used in an iterative fashion for accurate early detection of patients in the process of developing clinical PE and can also provide prognostic information related to the severity of the final diagnosis.

Serum sFlt1 and PlGF are protein analytes that have been widely reported to be dysregulated in PE. Generally, the sFlt1:PlGF ratio has shown better diagnostic performance compared to using

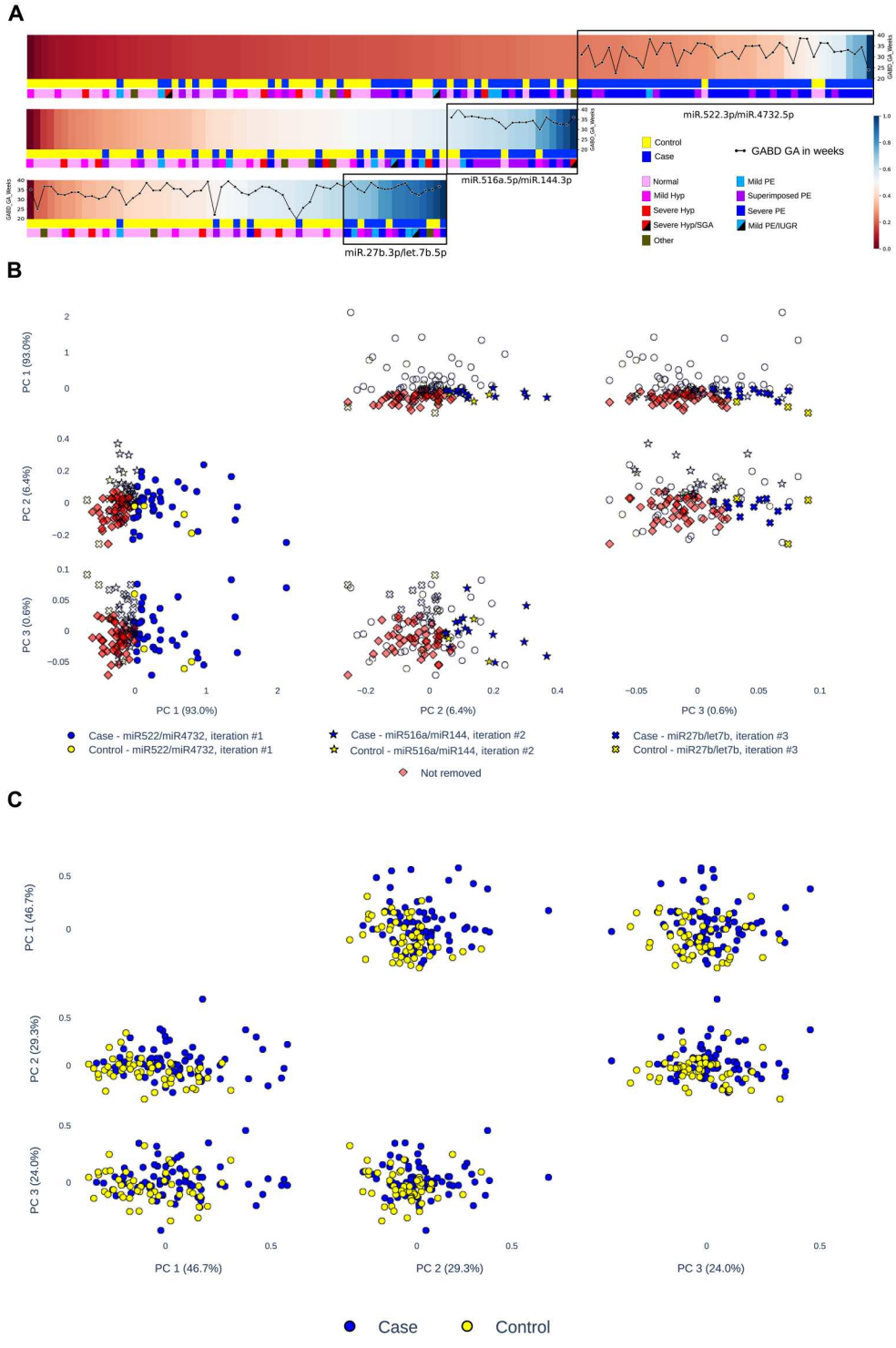


Fig. 7. Panel of three candidate bivariate biomarkers to discriminate between cases and controls. (**A**) Heatmaps showing normalized ratio of each of the three bivariates selected using iterative machine learning approach (see Materials and Methods) in the entire cohort (discovery and verification; $n = 82$ and $n = 41$, respectively). No patients were included in multiple classifications. Black boxes indicate samples selected after each iteration (bivariate). Colored bars at the bottom of each heatmap display diagnosis and detailed diagnosis of samples in heatmap. Line graph superimposed on top of heatmap displays the gestational age in weeks at blood draw. (**B**) Principal components analysis (first three principal components) of the three selected bivariate biomarkers. Blue (cases) and yellow (controls) dots show samples removed in first iteration (left column), blue (cases) and yellow (controls) stars show samples removed in second iteration (middle column), and blue (cases) and yellow (controls) x's show samples removed in third iteration (right column). Red diamonds are samples not removed during iteration, and white symbols show cases that were removed during a different iteration. (**C**) Principal components analysis (first three principal components) of three randomly selected bivariates from the top 1000 ranked bivariates. Blue dots show cases, and yellow dots show controls. PC, protein-to-creatinine; GABD, gestational age at blood draw; GA, gestational age; PE, preeclampsia.

either biomarker alone (*46*–*49*). We were able to acquire sFlt1 and PlGF Roche data for 70 of our patients (table S7). Currently, a sFlt1:PlGF ratio of 38 or above is used to predict diagnosis of PE within 4 weeks for women in whom the syndrome is suspected clinically (*11*). Using this cutoff, we calculated a sensitivity of 74.5% and a specificity of 91.3% for sFlt1:PlGF in our subjects. As described above, using our bivariate miRNA biomarkers, we achieved a sensitivity over 90% for both the discovery and validation cohort and a specificity of 79%. Restricting the analysis to only the 70 subjects who had sFlt1:PlGF ratio data, the sensitivity of our miRNA biomarkers was 89.3% and the specificity was 73.9%. Because of the multi-step process used for early diagnosis of PE using our bivariate miRNA biomarkers, it is not possible to calculate the sensitivity of these biomarkers at a fixed specificity. Instead, to enable

more direct comparison of the performance of our bivariate biomarker assay with that of the sFlt1:PlGF ratio, we adjusted the sFlt1:PlGF cutoff to match the specificity of our assay as closely as possible (sFlt1:PlGF 12.1, specificity of 78.3%), and obtained a sensitivity of 89.4%. This analysis suggests that the bivariate miRNA biomarker assay has a similar performance to the recently FDA-approved sFlt1:PlGF assay.

We then used the 70 participants that had both small RNA-seq and the sFlt1:PlGF measurements and explored whether combining the bivariate miRNA assay with the sFlt1:PlGF assay would result in improved performance compared to either assay alone. Inspecting the 48 cases that were designated as PE by the miRNA assay (42 of which were true positives and 6 of which were false positives), applying a cutoff of 12 for the sFlt1:PlGF assay would eliminate the 6 false positives (by correctly identifying them as normal) while retaining 39 of the true positives (with 3 being incorrectly reassigned from PE to normal). Inspecting the 22 cases that were designated as normal by the miRNA assay (17 of which were true negatives and 5 of which were false negatives), applying a cutoff of 32 for the sFlt1:PlGF assay would correctly identify 3 of the false negatives as PE, and retain 2 of the false negatives as normal, retain 15 of the true negatives, and incorrectly reassign 2 of the true negatives as PE. Overall, applying the miRNA bivariate biomarkers first, followed by the sFlt1:PlGF ratio, resulted in 42 true positives, 2 false positives, 21 true negatives, 5 false negatives, a sensitivity of 89.4%, specificity of 91.3%, PPV of 95.5%, and NPV of 80.8%. Alternatively, applying a secondary sFlt1:PlGF assessment only to the participants who were identified as positive by the bivariate miRNA assay results in 39 true positives, 0 false positives, 23 true negatives, and 8 false negatives, for a sensitivity of 82.9% and a specificity of 100%. This suggests that applying the bivariate miRNA biomarkers and the sFlt1:PlGF ratio sequentially can improve the sensitivity and specificity for early diagnosis of PE over either set of biomarkers alone.

DISCUSSION

PE is an important cause of fetal and maternal morbidity and mortality, for which there are currently no highly accurate methods for early diagnosis or prognostic assessment. Here, we used an unbiased transcriptomic approach to identify and verify univariate (single miRNAs) and bivariate (pairs of miRNAs) extracellular miRNA biomarkers for the early diagnosis of PE in discovery and verification cohorts of pregnant women being evaluated in a triage unit for signs and symptoms of PE. Consistent with our previous study focused on prediction of PE in an asymptomatic cohort of pregnant women (*45*), we were able to verify a larger number of bivariate biomarkers compared to univariate biomarkers, presumably due to the mutual normalizing effect of the miRNA pairs that accounted for the biological and technical variability inherent in small RNA-seq analysis of human biofluid samples. Moreover, we found that in addition to separating cases (patients who developed PE) from controls (those that did not), these ex-miRNA biomarkers could be used to distinguish among different categories of hypertensive disease in pregnancy.

Many previous studies in this area have focused on miRNA known to be enriched in the placenta compared to other organs. However, given that PE is associated with dysfunction of maternal organs, as well as the placenta, we used an unbiased comprehensive transcriptomic approach that enabled us to interrogate maternal serum miRNAs without regard to the source tissue. One hundred twenty-three patients were enrolled from an obstetrical triage unit, where they were being evaluated for signs and/or symptoms of PE and followed until delivery, to determine whether they were diagnosed with PE before delivery (cases) or not (controls). The resulting 71 cases and 52 controls were divided into matched discovery and verification cohorts. The cases included subjects with a range of manifestations of PE, and controls included both nonhypertensive subjects and subjects with chronic hypertension and mild gestational hypertension. Cases and controls were well matched in terms of clinical factors apart from race/ethnicity and differences secondary to known associations between PE and iatrogenic delivery and fetal growth restriction, including median GA at delivery, mean birthweight, SGA, and admission to the NICU.

We identified the top 100 univariate ex-miRNA biomarkers in the discovery cohort, re-ranked them using the verification cohort, and selected the top 10 for further examination. We found that performing clustering using these 10 univariate biomarkers not only separated cases and controls but also distinguished between detailed diagnoses and correlated with the interval between blood draw and PE diagnosis. The top 10 included *miR-144-3p*, *miR-1323*, *miR-518e-5p*, and *miR-516a-5p*, which have been reported in previous ex-miRNA studies to be associated with PE (*26*, *31*, *35*, *39*, *41*, *42*). The two univariate candidate PE biomarkers that were statistically significant in both the discovery and verification cohorts were *let-7b-5p* and *miR-423-5p*. Both ex-miRNA biomarkers were found at lower levels in our PE cases compared to controls. For *let-7b-5p*, our results are consistent with Gunel *et al.* (*28*), who also reported that let-7b-5p was present at lower levels in the plasma of PE cases compared to controls. However, *miR-423-5p* has been reported in previous studies to be up-regulated in maternal plasma in the first trimester of women who later developed early severe PE (*37*, *50*) and in women diagnosed with PE (*51*). Guo *et al.* (*51*) also reported that *miR-423-5p* was expressed at higher levels in the placentas of patients with PE compared to controls and that it inhibits trophoblast migration, invasion, and proliferation via *IGF2BP1*.

Our bivariate analysis yielded 110 bivariate ex-miRNA biomarkers that passed both discovery and verification criteria. The top bivariate ex-miRNA biomarkers by frequency and their appearance in previous reports are shown in data file S8. Two of our top ranked miRNA univariate biomarkers, *miR-423-5p* and *miR-20a-5p*, were also over-represented in our verified set of bivariate biomarkers, which also contained a disproportionate number of miRNAs from chromosome 19. As noted above, extracellular *miR-423-5p* has been reported to be more highly expressed in early pregnancy in patients who later develop PE by other groups (*37*, *50*). Our own group also identified this miRNA as the numerator of an early predictive bivariate biomarker for PE (*45*); we used the same convention in this previous study, where numerators were more highly expressed in PE. This raises the possibility that *miR-423-5p* may be expressed differently in the maternal circulation in presymptomatic and symptomatic PE.

miR-20a has been previously linked with PE and has been reported to inhibit proliferation, migration, and invasion of JEG3 choriocarcinoma cells by targeting *FOXA1* (*52*). As noted above, our verified bivariate biomarkers are enriched for miRNAs encoded in a region on chromosome 19 commonly referred to as the C19MC cluster. Aberrant expression of miRNAs in the

C19MC has been previously correlated with PE (*29*, *53*), although the three most highly enriched miRNAs among our verified bivariate biomarkers, *miR-512*, *miR-516*, and *miR-522*, have not been previously linked with PE. The C19MC has been found to be particularly highly expressed in the placenta and in human pluripotent stem cells (*54*) and has recently been reported to suppress genes critical for maintaining the epithelial cytotrophoblast stem cell phenotype and linked to the expression of several genes involved in cell migration and proliferation (*55*, *56*). It has been suggested that dysregulation of this miRNA cluster may result in impaired invasion associated with the shallow placentation of PE (*55*).

A markedly higher number of bivariate biomarkers passed in both discovery and verification phases of our study compared to univariate biomarkers. The bivariate biomarkers not only showed better separation of cases and controls but also were able to more clearly distinguish among subjects in phenotypic subgroups. We suggest that the superior performance of the bivariate biomarkers may be due to the internal normalization provided by the two miRNAs in each bivariate biomarker to account for technical and biological variability.

In post hoc analysis, we used a hierarchical, agglomerative clustering package using a weighted method and a custom Pearson correlation distance metric to cluster our verified bivariate biomarkers into 10 clusters, which not only displayed different patterns of expression across phenotypic subgroups but also showed enrichment of miRNAs encoded in similar genomic locations, suggesting that miRNAs transcribed from the same region of the genome are coherently regulated by the physiological differences between the phenotypic subgroups. When we analyzed the potential cell or tissue source of these bivariate biomarkers, we found that they largely originated from the placenta, liver, platelets, or RBCs. This was not unexpected, given the fact that dysfunction of several of these cells/tissues is associated with PE (placenta, liver, and platelets), and that they either are highly vascularized or are components of the maternal blood (and therefore are well sampled by serum, the biofluid used for this study). Consistent with a previous study reporting discovery and verification of early predictive ex-miRNAs for PE in asymptomatic gravidas from our group (*45*), this current study found that bivariate ex-miRNA biomarkers for early diagnosis of PE in symptomatic gravidas were largely composed of one miRNA from the placenta and one from a nonplacental source, suggesting that miRNAs expressed by nonplacental cells/tissues may serve to normalize the placental contributions.

To identify the smallest number of bivariate ex-miRNA biomarkers that could efficiently separate cases from controls, we applied a machine learning approach in an iterative fashion. This consisted of ranking the 110 bivariate biomarkers that passed discovery and verification using XGBoost and using the top-ranked biomarker to identify high-probability PE cases, which were then set aside. XGBoost was then reapplied to the remaining bivariate biomarkers and the remaining cases and controls, two more times. Overall, this iterative process was able to identify over 90% of cases with a false-positive rate of less than 10% at each iteration and achieved a PPV of 55% at a sensitivity of 93% and specificity of 79% (+LR = 4.43, −LR = 0.09) given a prevalence of PE of 57.7% in our study. In the first round, the identified cases were predominantly those that developed PE with severe features, as well as a smaller number of superimposed PE cases. The second round identified most of the remaining superimposed PE cases and our one severe hypertension/SGA case. The third round identified all but one remaining PE case, which was a mild PE case. The controls misidentified as PE using this process were all patients diagnosed with mild hypertension or gestational proteinuria. No control patients with severe hypertension were misclassified as PE. These results suggest that some patients with mild hypertension or gestational proteinuria may share some physiological features with PE that lead to similar alterations in ex-miRNA expression, whereas severe hypertension may actually be physiologically more distinct from PE. Using a separate cohort of 18 patients, we found all but one PE case and achieved a sensitivity of 91% and specificity of 57%.

Our approach revealed both univariate and bivariate miRNA biomarkers, mapped out a method to predict the development of PE using three bivariate biomarkers in an iterative manner, and validated this approach in a separate patient cohort. The candidate biomarkers we have found in this cohort of patients will now need to be confirmed on a larger multicenter independent cohort. Future validation studies are required to establish the clinical utility of this approach for early diagnosis and prognosis of PE in the obstetrical triage setting and to better define performance limits of the approach regarding factors such as the interval from blood draw to PE diagnosis. Given our additional finding that combining our candidate miRNA biomarkers with the sFlt1:PlGF ratio results in improved sensitivity and specificity over either the miRNA or sFlt1:PlGF biomarkers alone, these future validation studies should include an evaluation of the utility and performance of this combined approach. Moreover, future in vitro studies to uncover the target genes regulated by our candidate miRNA biomarkers may have utility in providing insights into the mechanistic basis for different PE subtypes. We believe that validation and clinical application of our candidate biomarkers will allow for better clinical resource allocation, prevent low-risk patients from unnecessary admission and procedures, and increase understanding of their roles in PE disease pathogenesis, which may help develop therapies for patients at high risk for the development of PE.

MATERIALS AND METHODS

Study design

Patients were consented under an Institutional Review Board (IRB) protocol approved by the Human Research Protections Program at UC San Diego. Serum samples and the RNA samples isolated from them were labeled with study identifiers with no personal identifiable information. Each subject signed a HIPAA (Health Insurance Portability and Accountability Act) release and agreed to have demographic and clinical data collected from the electronic medical record (EPIC) throughout the course of pregnancy and postpartum. The primary objective of the study was to use an unbiased transcriptomic approach to discover ex-miRNA biomarkers for the early diagnosis of PE in a cohort of patients presenting for evaluation for possible PE. A secondary objective was to find ex-miRNA biomarkers that predict the severity of PE in those who developed this complication. Our sample size was determined by the number of patients consented during a given time frame, and data inclusion and exclusion criteria are detailed below and in the text. Patients were randomly assigned to the discovery and verification groups, and investigators performing the laboratory analysis were blinded to patient's diagnosis.

Recruitment and enrollment

Patients being evaluated in the UC San Diego OB Triage unit, between 20^0 and 40^6 weeks (20 weeks 0 days and 40 weeks 6 days) GA, presenting with suspected PE (e.g., complaint of headache, visual disturbances, epigastric pain, hypertension, proteinuria, or fetal growth restriction), were approached and screened for enrollment. Patients who received a diagnosis of PE before delivery were classified as cases, while all patients who did not develop PE were classified as controls. All diagnoses were determined via retrospective review of the medical charts at least 8 weeks after delivery.

Inclusion criteria were as follows: pregnant women age $\geq$ 18, prenatal care at UC San Diego Prenatal Care Clinics, singleton pregnancy, GA between 20^0 and 40^6 weeks GA, and undergoing evaluation for suspected PE.

Exclusion criteria were as follows: previous admission during the current pregnancy for suspected PE, previously diagnosed during the current pregnancy with severe preterm PE, plan for immediate delivery, and admission to the hospital for other indications, with post-admission development of signs or symptoms requiring evaluation for PE

Adjudications of pregnancy outcomes

The clinical outcome of each pregnancy was adjudicated by two obstetrician-gynecologists, at least one of which was board-certified in maternal fetal medicine.

Sample collection

After informed consent was obtained, 10 ml of blood was collected by peripheral venipuncture, allowed to clot in an upright position for at least 20 min at room temperature, and then centrifuged at 2000*g* for 10 min. The serum was removed and placed into aliquot tubes each containing 0.5 ml of serum and stored at −80°C. Sample processing was completed within 2 hours of blood collection. Extracellular RNA (exRNA) isolation and analysis was performed retrospectively, after delivery. Thus, exRNA levels were not known by the subject, physician, or study team during the pregnancy and had no impact upon the evaluation and management of the subject. One hundred thirty-one subjects were recruited and donated samples for the initial discovery and verification groups. An additional 18 subjects were recruited and donated samples to verify candidate bivariate biomarkers.

Laboratory analysis

exRNA was isolated from each serum sample using the Plasma/Serum Circulating and Exosomal RNA Purification Kit (Slurry Format) (Norgen Biotek Corp., Ontario, Canada). Quality control of isolated RNA was performed using the Agilent RNA 6000 Pico Kit (Agilent, Santa Clara, CA). Small RNA-seq libraries were constructed using the NEB-Next Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs Inc., Ipswich, MA). The libraries were then cleaned using the DNA Clean & Concentrator-5 Kit (Zymo Research, Irvine, CA), and quality control of the libraries was performed using the Agilent High Sensitivity DNA Kit (Agilent, Santa Clara, CA). Equal volumes of the libraries were pooled, size-selected for products that were 120 to 135 base pairs in length using a Pippin Prep with a 3% agarose gel cassette (Sage Science, Beverly, MA), and run on a MiSeq instrument at the UC San Diego Institute for Genomic Medicine (IGM) Genomics Core using the MiSeq Nano Reagent Kit (Illumina, San Diego, CA). Samples that produced adequate numbers of miRNA read counts were then rebalanced to produce similar numbers of miRNA reads and sequenced on a HiSeq 4000 instrument to produce 1 × 75 bp reads (Illumina, San Diego, CA) at the UC San Diego IGM Genomics Core.

Sequencing data analysis

Small RNA-seq data were trimmed and mapped to known human sequences using the exceRpt Small RNA-seq Pipeline Workflow implemented in the Genboree Workbench (*44*). Each sample was evaluated for miRNA read depth and miRNA complexity. Samples that exhibited low miRNA read depth (<500,000 miRNA reads) or low miRNA complexity (<300 different miRNA species) were excluded from analysis. miRNAs were filtered such that miRNAs with at least 10 raw reads in at least 50% of cases or controls were retained, resulting in 267 pass-filter miRNAs. The filtered and scaled sequencing data were normalized by global scaling and visualized using the Qlucore Omics Explorer Software (Qlucore, Lund, Sweden) and evaluated using the statistical analysis methods described below. miRNA tissue contributions were determined using the recently reported fractional contribution of each cell/tissue type to each miRNA (*45*). For each miRNA, the tissue with the maximum contribution and any tissue within 10% of the maximum were considered to be a contributing tissue source.

PEER

Small RNA-seq data were filtered such that miRNAs were required to have reads detected in at least 25% of the patients. This resulted in an input dataset of 619 miRNAs. Cases and controls were divided among discovery (training) and verification groups aiming to maintain a similar GA range in both groups. The raw counts for the remaining miRNAs are provided in data file S7. Starting with the discovery group, read counts were $\log_2$-transformed and loess-normalized. The PEER package (v.1.0) (*57*) was then run to model the effect of our cases and controls. The PEER process performance was assessed using principal components analysis and Kruskal-Wallis test. Using the PEER processed data, a *P* value comparing the cases and controls was calculated for each miRNA using a generalized linear model and chi-square test. AUCs were then generated for each miRNA with the pROC package, using the Delong and bootstrap methods to establish confidence intervals and a *t* test to generate *P* values (*58*). As previously published (*45*), bivariate datasets were created with all possible miRNA ratios and ranked by an inverse rank sum using 1000 bootstraps. At each iteration, AUCs were calculated in a similar manner to the univariate data, as well as the squared correlation between the miRNA ratio and the binary case/control column, and the mean difference between cases and controls. Five ranks were derived from the resulting 1000 iterations for each miRNA ratio: (i) the mean of the AUCs, (ii) the lower 25% quantile of the AUCs, (iii) the mean of the squared correlation, (iv) the upper 75% quantile of the AUCs, and (v) the *P* value calculated using a *t* test. Each rank was then inverted and summed for each miRNA ratio to obtain a final ranking (data file S9). The same procedure was used for processing the verification group. Analysis was performed using R 3.4.1. Clustering of bivariate biomarkers was done using the scipy hierarchical, agglomerative clustering package using a weighted method and a custom Pearson correlation distance metric. The median, *z*-scaled bivariate ratios were used as input into the clustering algorithm. Enrichment analysis of the 110

bivariate biomarkers was done using the CDF of the hypergeometric distribution using the top 10,000 bivariate biomarkers as the initial population size.

The discovery and verification datasets were processed separately. Ranked values were combined with the calculated statistics for each of the miRNA ratios. The data were then filtered for only miRNA ratios with an AUC of >0.5 and then sorted by their calculated rank. The top 1000 ranked miRNA ratios were then assessed in the verification dataset. We considered any miRNA ratio that had a lower 25% quantile AUC over 0.7 in the verification dataset "verified" (110 miRNA ratios). To assess the performance of our method, we performed principal components analysis on both the set of verified miRNA ratios and 110 randomly selected ratios. The ratio values used in the principal components analysis were obtained by combining the discovery and verification datasets before performing normalization. Combined discovery and verification miRNA ratio ranking values, the raw AUCs, the squared correlations, and *P* values the ranks were based on, the mean and median of both the cases and controls for the separate miRNAs in the ratio, and the normalized ratio values from the combined dataset for the top 1000 miRNA ratios in the discovery dataset are provided in data file S10.

The univariate discovery and verification dataset were created separately using the AUCs and chi-square *P* values generated for each of the 619 miRNAs and then ranked using the chi-square values. The top 100 ranked miRNAs from the univariate discovery dataset were extracted from the univariate verification dataset and ranked based on the chi-square *P* value. To assess the performance of our method, we performed principal components analysis using all patient samples on all 619 miRNAs, the top 100 discovery miRNAs, and the top 10 miRNAs based on the verification dataset ranking. Normalized univariate data combined with calculated AUCs, squared correlation, chi-square *P* values, and *t* test *P* values for both the discovery and verification groups are provided in data file S11.

Candidate bivariate marker selection

Feature selection of the 110 verified bivariate ratios was performed using an XGBClassifier model from the python package XGBoost (v. 1.4.0). The test size was set at 0.15, random state equaled 42, and feature importance values were obtained from the coefficients. The bivariate found to have the largest feature importance was used to sort the 110 bivariates, and all samples higher than 90% of the control samples along with the bivariate were removed for the second iteration of the XGBClassifier. The process was repeated for a total of three iterations/bivariates. Principal components analysis of all the samples using the three highest bivariates was performed using principal components analysis (PCA) from sklearn (v. 0.23.2) and plotted using plotly (v. 4.7.0). Sensitivity was calculated as True Positives/(True Positives + False Negatives), specificity was calculated as True Negatives/(True Negatives + False Positives), PPV was calculated as PPV = True Positives/(True Positives + False Positives), and NPV was calculated as NPV = True Negatives/(True Negatives + False Negatives).

Statistical analysis

Clinical data were analyzed using Student's *t* test, Mann-Whitney *U* test, and Fisher's exact test with SPSS version 25. The testing level and adjustments are detailed in Results and Materials and Methods each time a statistical test was used. Hierarchical clustering was performed and displayed using the Qlucore Omics Explorer Software (Qlucore, Lund, Sweden) to visualize candidate miRNAs identified using the PEER process. Clustering and enrichment of the 110 bivariate biomarkers was done using the CDF and is detailed above.

Supplementary Materials

This PDF file includes:
Figs. S1 to S3
Legends for data files S1 to S11

Other Supplementary Material for this manuscript includes the following:
Data files S1 to S11

REFERENCES AND NOTES

1. R. B. Ness, J. M. Roberts, Heterogeneous causes constituting the single syndrome of preeclampsia: A hypothesis and its implications. *Am. J. Obstet. Gynecol.* **175**, 1365–1370 (1996).
2. B. M. Sibai, Diagnosis and management of gestational hypertension and preeclampsia. *Obstet. Gynecol.* **102**, 181–192 (2003).
3. C. W. Redman, I. L. Sargent, Latest advances in understanding preeclampsia. *Science* **308**, 1592–1594 (2005).
4. Y. Zhou, C. H. Damsky, S. J. Fisher, Preeclampsia is associated with failure of human cytotrophoblasts to mimic a vascular adhesion phenotype. One cause of defective endovascular invasion in this syndrome? *J. Clin. Invest.* **99**, 2152–2164 (1997).
5. G. J. Burton, C. W. Redman, J. M. Roberts, A. Moffett, Pre-eclampsia: Pathophysiology and clinical implications. *BMJ* **366**, l2381 (2019).
6. K. Webster, S. Fishburn, M. Maresh, S. C. Findlay, L. C. Chappell, Diagnosis and management of hypertension in pregnancy: Summary of updated NICE guidance. *BMJ* **366**, l5119 (2019).
7. I. Herraiz, E. Llurba, S. Verlohren, A. Galindo; Spanish Group for the Study of Angiogenic Markers in Preeclampsia, Update on the diagnosis and prognosis of preeclampsia with the aid of the sFlt-1/ PlGF ratio in singleton pregnancies. *Fetal Diagn. Ther.* **43**, 81–89 (2018).
8. G. K. Frampton, J. Jones, M. Rose, L. Payne, Placental growth factor (alone or in combination with soluble fms-like tyrosine kinase 1) as an aid to the assessment of women with suspected pre-eclampsia: Systematic review and economic analysis. *Health Technol. Assess.* **20**, 1–160 (2016).
9. C. Tardif, E. Dumontet, H. Caillon, E. Misbert, V. Dochez, D. Masson, N. Winer, Angiogenic factors sFlt-1 and PlGF in preeclampsia: Prediction of risk and prognosis in a high-risk obstetric population. *J. Gynecol. Obstet. Hum. Reprod.* **47**, 17–21 (2018).
10. I. Dragan, T. Georgiou, N. Prodan, R. Akolekar, K. H. Nicolaides, Screening for pre-eclampsia using sFlt-1/PlGF ratio cut-off of 38 at 30–37 weeks' gestation. *Ultrasound Obstet. Gynecol.* **49**, 73–77 (2017).
11. H. Zeisler, E. Llurba, F. Chantraine, M. Vatish, A. C. Staff, M. Sennström, M. Olovsson, S. P. Brennecke, H. Stepan, D. Allegranza, P. Dilba, M. Schoedl, M. Hund, S. Verlohren, Predictive value of the sFlt-1:PlGF ratio in women with suspected preeclampsia. *N. Engl. J. Med.* **374**, 13–22 (2016).
12. R. Thadhani, E. Lemoine, S. Rana, M. M. Costantine, V. F. Calsavara, K. Boggess, B. J. Wylie, T. A. Moore Simas, J. M. Louis, J. Espinoza, S. L. Gaw, A. Murtha, S. Wiegand, Y. Gollin, D. Singh, R. M. Silver, D. E. Durie, B. Panda, E. R. Norwitz, I. Burd, B. Plunkett, R. K. Scott, A. Gaden, M. Bautista, Y. Chang, M. A. Diniz, S. A. Karumanchi, S. Kilpatrick, Circulating angiogenic factor levels in hypertensive disorders of pregnancy. *NEJM Evidence* **1**, EVIDoa2200161 (2022).
13. H. Duan, G. Zhao, B. Xu, S. Hu, J. Li, Maternal serum PLGF, PAPPA, β-hCG and AFP levels in early second trimester as predictors of preeclampsia. *Clin. Lab.* **63**, 921–925 (2017).
14. C. Birdir, L. Droste, L. Fox, M. Frank, J. Fryze, A. Enekwe, A. Köninger, R. Kimmig, B. Schmidt, A. Gellhaus, Predictive value of sFlt-1, PlGF, sFlt-1/PlGF ratio and PAPPA for late-onset preeclampsia and IUGR between 32 and 37 weeks of pregnancy. *Pregnancy Hypertens.* **12**, 124–128 (2018).
15. M. B. A. Gannoun, S. Bourrelly, N. Raguema, H. Zitouni, E. Nouvellon, W. Maleh, A. B. Chemili, R. Elfeleh, W. Almawi, T. Mahjoub, J.-C. Gris, Placental growth factor and vascular endothelial growth factor serum levels in Tunisian Arab women with suspected preeclampsia. *Cytokine* **79**, 1–6 (2016).

16. G. E. Osanyin, K. S. Okunade, A. A. Oluwole, Association between serum CA125 levels in preeclampsia and its severity among women in Lagos, South-West Nigeria. *Hypertens. Pregnancy* **37**, 93–97 (2018).
17. D. Morano, A. Rolfo, V. Tisato, V. Tisato, A. Farina, E. Rimondi, G. Scutiero, P. Greco, G. Bonaccorsi, T. Todros, Lower maternal serum tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) levels in early preeclampsia. A retrospective study. *Pregnancy Hypertens.* **12**, 1–5 (2018).
18. S. Liao, M. H. Vickers, R. S. Taylor, B. Jones, M. Fraser, L. M. E. McCowan, P. N. Baker, Maternal serum IGF-1, IGFBP-1 and 3, and placental growth hormone at 20 weeks' gestation in pregnancies complicated by preeclampsia. *Pregnancy Hypertens.* **10**, 149–154 (2017).
19. Y. Çekmez, Ş. Garip, İ. Ulu, S. Gülşen, E. T. Haberal, Y. Olgac, E. E. Yoğurtçuoğlu, S. B. Türkmen, F. T. Aksoy, G. Kıran, Maternal serum Netrin-1 levels as a new biomarker of preeclampsia. *J. Matern. Fetal Neonatal Med.* **30**, 1072–1074 (2017).
20. M. F. Siddiqui, P. Nandi, G. V. Girish, K. Nygard, G. Eastabrook, B. de Vrijer, V. K. M. Han, P. K. Lala, Decorin over-expression by decidual cells in preeclampsia: A potential blood biomarker. *Am. J. Obstet. Gynecol.* **215**, 361.e1–361.e15 (2016).
21. F. P. McCarthy, A. Doyle, A. S. Khashan, L. C. Kenny, Altered maternal plasma glycogen phosphorylase isoenzyme BB as a biomarker for preeclampsia and small for gestational age. *Reprod. Sci.* **23**, 738–747 (2016).
22. X. Luo, X. Li, Long non-coding RNAs Serve as diagnostic biomarkers of preeclampsia and modulate migration and invasiveness of trophoblast cells. *Med. Sci. Monit.* **24**, 84–91 (2018).
23. M. Y. Tan, A. Syngelaki, L. C. Poon, D. L. Rolnik, N. O'Gorman, J. L. Delgado, R. Akolekar, L. Konstantinidou, M. Tsavdaridou, S. Galeva, U. Ajdacka, F. S. Molina, N. Persico, J. C. Jani, W. Plasencia, E. Greco, G. Papaioannou, A. Wright, D. Wright, K. H. Nicolaides, Screening for pre-eclampsia by maternal factors and biomarkers at 11-13 weeks' gestation. *Ultrasound Obstet. Gynecol.* **52**, 186–195 (2018).
24. A. Farina, C. Zucchini, A. Sekizawa, Y. Purwosunu, P. de Sanctis, G. Santarsiero, N. Rizzo, D. Morano, T. Okai, Performance of messenger RNAs circulating in maternal blood in the prediction of preeclampsia at 10-14 weeks. *Am. J. Obstet. Gynecol.* **203**, 575.e1–575.e7 (2010).
25. M. Rasmussen, M. Reddy, R. Nolan, J. Camunas-Soler, A. Khodursky, N. M. Scheller, D. E. Cantonwine, L. Engelbrechtsen, J. D. Mi, A. Dutta, T. Brundage, F. Siddiqui, M. Thao, E. P. S. Gee, J. La, C. Baruch-Gravett, M. K. Santillan, S. Deb, S. M. Ame, S. M. Ali, M. Adkins, M. A. DePristo, M. Lee, E. Namsaraev, D. J. Gybel-Brask, L. Skibsted, J. A. Litch, D. A. Santillan, S. Sazawal, R. M. Tribe, J. M. Roberts, M. Jain, E. Høgdall, C. Holzman, S. R. Quake, M. A. Elovitz, T. F. McElrath, RNA profiles reveal signatures of future health and disease in pregnancy. *Nature* **601**, 422–427 (2022).
26. L. Wu, H. Zhou, H. Lin, J. Qi, C. Zhu, Z. Gao, H. Wang, Circulating microRNAs are elevated in plasma from severe preeclamptic pregnancies. *Reproduction* **143**, 389–397 (2012).
27. L. Gan, Z. Liu, M. Wei, Y. Chen, X. Yang, L. Chen, X. Xiao, MIR-210 and miR-155 as potential diagnostic markers for pre-eclampsia pregnancies. *Medicine* **96**, e7515 (2017).
28. T. Gunel, M. K. Hosseini, E. Gumusoglu, H. I. Kisakesen, A. Benian, K. Aydinli, Expression profiling of maternal plasma and placenta microRNAs in preeclamptic pregnancies by microarray technology. *Placenta* **52**, 77–85 (2017).
29. I. Hromadnikova, K. Kotlabova, K. Ivankova, L. Krofta, First trimester screening of circulating C19MC microRNAs and the evaluation of their potential to predict the onset of pre-eclampsia and IUGR. *PLOS ONE* **12**, e0171756 (2017).
30. I. Hromadnikova, K. Kotlabova, M. Ondrackova, A. Kestlerova, V. Novotna, L. Hympanova, J. Doucha, L. Krofta, Circulating C19MC micrornas in preeclampsia, gestational hypertension, and fetal growth restriction. *Mediators Inflamm.* **2013**, 186041 (2013).
31. D. S. Jairajpuri, Z. H. Malalla, N. Mahmood, W. Y. Almawi, Circulating microRNA expression as predictor of preeclampsia and its severity. *Gene* **627**, 543–548 (2017).
32. H. Li, Q. Ge, L. Guo, Z. Lu, Maternal plasma miRNAs expression in preeclamptic pregnancies. *Biomed. Res. Int.* **2013**, 970265 (2013).
33. A. Luque, A. Farwati, F. Crovetto, F. Crispi, F. Figueras, E. Gratacós, J. M. Aran, Usefulness of circulating microRNAs for the prediction of early preeclampsia at first-trimester of pregnancy. *Sci. Rep.* **4**, 4882 (2014).
34. M. L. Martinez-Fierro, I. Garza-Veloz, C. Gutierrez-Arteaga, I. Delgado-Enciso, O. Y. Barbosa-Cisneros, V. Flores-Morales, G. P. Hernandez-Delgadillo, M. R. Rocha-Pizaña, I. P. Rodriguez-Sanchez, J. I. Badillo-Almaraz, J. M. Ortiz-Rodriguez, R. Castañeda-Miranda, L. O. Solis-Sanchez, Y. Ortiz-Castro, Circulating levels of specific members of chromosome 19 microRNA cluster are associated with preeclampsia development. *Arch. Gynecol. Obstet.* **297**, 365–371 (2018).
35. K. Miura, A. Higashijima, Y. Murakami, O. Tsukamoto, Y. Hasegawa, S. Abe, N. Fuchi, S. Miura, M. Kaneuchi, H. Masuzaki, Circulating chromosome 19 miRNA cluster microRNAs in pregnant women with severe pre-eclampsia. *J. Obstet. Gynaecol. Res.* **41**, 1526–1532 (2015).
36. T. M. K. Motawi, D. Sabry, N. W. Maurice, S. M. Rizk, Role of mesenchymal stem cells exosomes derived microRNAs; miR-136, miR-494 and miR-495 in pre-eclampsia diagnosis and evaluation. *Arch. Biochem. Biophys.* **659**, 13–21 (2018).
37. C. Salomon, G. E. Rice, Role of exosomes in placental homeostasis and pregnancy disorders. *Prog. Mol. Biol. Transl. Sci.* **145**, 163–179 (2017).
38. J. Stubert, D. Koczan, D.-U. Richter, M. Dieterich, B. Ziems, H.-J. Thiesen, B. Gerber, T. Reimer, miRNA expression profiles determined in maternal sera of patients with HELLP syndrome. *Hypertens. Pregnancy* **33**, 215–235 (2014).
39. B. Ura, G. Feriotto, L. Monasta, S. Bilel, M. Zweyer, C. Celeghini, Potential role of circulating microRNAs as early markers of preeclampsia. *Taiwan. J. Obstet. Gynecol.* **53**, 232–234 (2014).
40. P. Xu, Y. Zhao, M. Liu, Y. Wang, H. Wang, Y.-X. Li, X. Zhu, Y. Yao, H. Wang, J. Qiao, L. Ji, Y.-L. Wang, Variations of microRNAs in human placentas and plasma from preeclamptic pregnancy. *Hypertension* **63**, 1276–1284 (2014).
41. Q. Yang, J. Lu, S. Wang, H. Li, Q. Ge, Z. Lu, Application of next-generation sequencing technology to profile the circulating microRNAs in the serum of preeclampsia versus normal pregnant women. *Clin. Chim. Acta* **412**, 2167–2173 (2011).
42. S. Yang, H. Li, Q. Ge, L. Guo, F. Chen, Deregulated microRNA species in the plasma and placenta of patients with preeclampsia. *Mol. Med. Rep.* **12**, 527–534 (2015).
43. L. Yoffe, A. Gilam, O. Yaron, A. Polsky, L. Farberov, A. Syngelaki, K. Nicolaides, M. Hod, N. Shomron, Early detection of preeclampsia using circulating small non-coding RNA. *Sci. Rep.* **8**, 3401 (2018).
44. J. Rozowsky, R. R. Kitchen, J. J. Park, T. R. Galeev, J. Diao, J. Warrell, W. Thistlethwaite, S. L. Subramanian, A. Milosavljevic, M. Gerstein, exceRpt: A comprehensive analytic platform for extracellular RNA profiling. *Cell Syst.* **8**, 352–357.e3 (2019).
45. S. Srinivasan, R. Treacy, T. Herrero, R. Olsen, T. R. Leonardo, X. Zhang, P. De Hoff, C. To, L. G. Poling, A. Fernando, S. Leon-Garcia, K. Knepper, V. Tran, M. Meads, J. Tasarz, A. Vuppala, S. Park, C. D. Laurent, T. Bui, P. S. Cheah, R. T. Overcash, G. A. Ramos, H. Roeder, I. Ghiran, M. Parast; PAPR Study Consortium, X. O. Breakefield, A. J. Lueth, S. R. Rust, M. T. Dufford, A. C. Fox, D. E. Hickok, J. Burchard, J. J. Boniface, L. C. Laurent, Discovery and verification of extracellular miRNA biomarkers for non-invasive prediction of pre-eclampsia in asymptomatic women. *Cell Rep. Med.* **1**, 100013 (2020).
46. S. Verlohren, A. Galindo, D. Schlembach, H. Zeisler, I. Herraiz, M. G. Moertl, J. Pape, J. W. Dudenhausen, B. Denk, H. Stepan, An automated method for the determination of the sFlt-1/PlGF ratio in the assessment of preeclampsia. *Am. J. Obstet. Gynecol.* **202**, 161.e1–161.e11 (2010).
47. S. Verlohren, I. Herraiz, O. Lapaire, D. Schlembach, M. Moertl, H. Zeisler, P. Calda, W. Holzgreve, A. Galindo, T. Engels, B. Denk, H. Stepan, The sFlt-1/PlGF ratio in different types of hypertensive pregnancy disorders and its prognostic potential in preeclamptic patients. *Am. J. Obstet. Gynecol.* **206**, 58.e1–58.e8 (2012).
48. S. Verlohren, I. Herraiz, O. Lapaire, D. Schlembach, H. Zeisler, P. Calda, J. Sabria, F. Markfeld-Erol, A. Galindo, K. Schoofs, B. Denk, H. Stepan, New gestational phase-specific cutoff values for the use of the soluble fms-like tyrosine kinase-1/placental growth factor ratio as a diagnostic test for preeclampsia. *Hypertension* **63**, 346–352 (2014).
49. I. Álvarez-Fernández, B. Prieto, V. Rodríguez, Y. Ruano, A. I. Escudero, F. V. Álvarez, New biomarkers in diagnosis of early onset preeclampsia and imminent delivery prognosis. *Clin. Chem. Lab. Med.* **52**, 1159–1168 (2014).
50. A. V. Timofeeva, V. A. Gusar, N. E. Kan, K. N. Prozorovskaya, A. O. Karapetyan, O. R. Bayev, V. V. Chagovets, S. F. Kliver, D. Y. Iakovishina, V. E. Frankevich, G. T. Sukhikh, Corrigendum to "Identification of potential early biomarkers of preeclampsia" [Placenta (2018) 61-71]. *Placenta* **63**, 61 (2018).
51. L. Guo, Y. Liu, Y. Guo, Y. Yang, B. Chen, MicroRNA-423-5p inhibits the progression of trophoblast cells via targeting IGF2BP1. *Placenta* **74**, 1–8 (2018).
52. Y. Wang, Y. Zhang, H. Wang, J. Wang, Y. Zhang, Y. Wang, Z. Pan, S. Luo, Aberrantly upregulated miR-20a in pre-eclampsic placenta compromised the proliferative and invasive behaviors of trophoblast cells by targeting forkhead box protein A1. *Int. J. Biol. Sci.* **10**, 973–982 (2014).
53. I. Hromadnikova, K. Kotlabova, M. Ondrackova, P. Pirkova, A. Kestlerova, V. Novotna, L. Hympanova, L. Krofta, Expression profile of C19MC microRNAs in placental tissue in pregnancy-related complications. *DNA Cell Biol.* **34**, 437–457 (2014).
54. I. Bentwich, A. Avniel, Y. Karov, R. Aharonov, S. Gilad, O. Barad, A. Barzilai, P. Einat, U. Einav, E. Meiri, E. Sharon, Y. Spector, Z. Bentwich, Identification of hundreds of conserved and nonconserved human microRNAs. *Nat. Genet.* **37**, 766–770 (2005).
55. E. F. Mong, Y. Yang, K. M. Akat, J. Canfield, J. Van Wye, J. Lockhart, J. C. M. Tsibris, F. Schatz, C. J. Lockwood, T. Tuschl, U. A. Kayisli, H. Totary-Jain, Chromosome 19 microRNA cluster enhances cell reprogramming by inhibiting epithelial-to-mesenchymal transition. *Sci. Rep.* **10**, 3029 (2020).

56. J.-F. Mouillet, J. Goff, E. Sadovsky, H. Sun, T. Parks, T. Chu, Y. Sadovsky, Transgenic expression of human C19MC miRNAs impacts placental morphogenesis. *Placenta* **101**, 208–214 (2020).
57. O. Stegle, L. Parts, M. Piipari, J. Winn, R. Durbin, Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat. Protoc.* **7**, 500–507 (2012).
58. X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, M. Müller, pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77 (2011).

Acknowledgments: We thank M. Jacobs and C. Luevano-Diaz for their help with identification of cases and controls for the validation cohort, and abstraction of metadata, from the Center for Perinatal Discovery's Perinatal Biorepository and Database. **Funding:** This work was supported by National Institutes of Health grants UH3TR000906, T32HD007203 (S.S.), T32GM008806 (R.M.), and R00HD096125 (P.P.); Eunice Kennedy Shriver National Institute of Child Health and Human Development grant NIH R00HD096125 (P.P.); National Institutes of Health grant P30CA023100 (MiSeq and HiSeq next-generation sequencing was conducted at the IGM Genomics Center, University of California, San Diego, La Jolla, CA); and Preeclampsia Foundation Canada Vision Grant 2021 (P.P.). Data management, storage, and analysis were performed using the Extreme Science and Engineering Discovery Environment (XSEDE) Expanse at the San Diego Supercomputing Center through allocation MCB140074, as well as the Genboree exceRpt Small RNA-seq pipeline. **Author contributions:** Conceptualization: R.M., L.P., and L.C.L. Methodology: R.M., L.P., and L.C.L. Investigation: R.M., L.P., S.S., C.M.-K., A.A., K.Z.-R., C.T., A.H., M.Mo., K.V., A.M., V.T., M.Me., L.L.-S., M.H., G.A.R., P.D., M.M.P., P.P., L.C.L., and H.R. Visualization: R.M. and L.P. Funding acquisition: L.C.L. Supervision: P.D., M.M.P., P.P., and L.C.L. Writing—original draft: R.M., L.P., and L.C.L. Writing—review and editing: R.M., L.P., S.S., C.M.-K., A.A., K.Z.-R., C.T., A.H., M.Mo., K.V., A.M., V.T., M.Me., L.L.-S., M.H., G.A.R., P.D., M.M.P., P.P., L.C.L., and H.R. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Data are available in dbGaP under the name "Extracellular microRNA Biomarkers for Diagnostic and Prognostic Assessment of Preeclampsia at Triage" under the accession number: phs003169.v1.

Submitted 18 January 2023
Accepted 17 November 2023
Published 20 December 2023
10.1126/sciadv.adg7545