

## ORIGINAL ARTICLE

# Mining DNA methylation alterations towards a classification of placental pathologies

Samantha L. Wilson<sup>1,2</sup>, Katherine Leavey<sup>3</sup>, Brian J. Cox<sup>3,4</sup> and Wendy P. Robinson<sup>1,2,\*</sup>

<sup>1</sup>BC Children's Hospital Research Institute, Vancouver, BC V5Z 4H4, Canada, <sup>2</sup>Department of Medical Genetics, University of British Columbia, Vancouver, BC V6H 3N1, Canada, <sup>3</sup>Department of Physiology, University of Toronto, Toronto, ON M5S 1A8, Canada and <sup>4</sup>Department of Obstetrics and Gynecology, University of Toronto, Toronto, ON M5G 1E2, Canada

\*To whom correspondence should be addressed at: BC Children's Hospital Research Institute, 950 W 28<sup>th</sup> Ave, Vancouver, BC V5Z 4H4, Canada. Tel: 6048753229; Fax: 6048753120; Email: wrobinson@bcchr.ca

## Abstract

Placental health is a key component to a successful pregnancy. Placental insufficiency (PI), inadequate nutrient delivery to the fetus, is associated with preeclampsia (PE), a maternal hypertensive disorder, and intrauterine growth restriction (IUGR), pathologically poor fetal growth. PI is more common in early-onset PE (EOPE) than late-onset PE (LOPE). However, the relationship between these disorders remains unclear. While DNA methylation (DNAm) alterations have been identified in PE and IUGR, these entities can overlap and few studies have analysed them separately. This study aims to utilize DNAm profiling to better understand the underlying placental variation associated with PE and IUGR. Placental samples from a discovery (43 controls, 22 EOPE, 18 LOPE, 11 IUGR) and validation cohort (15 controls, 22 EOPE, 11 LOPE) were evaluated using the Illumina HumanMethylation450 array. To account for gestational age (GA) effects, EOPE samples were compared with pre-term births of varying etiologies (GA < 37 weeks). LOPE and IUGR were compared with term controls (GA > 37 weeks). While 1703 sites were differentially methylated (DM) (FDR < 0.05,  $\Delta\beta > 0.1$ ) in EOPE, few changes were associated with LOPE (N = 5), or IUGR (N = 0). Of the 1703 EOPE sites, 599 validated in the second cohort. Using these 599 sites, both cohorts clustered into three distinct groups. Interestingly, LOPE samples diagnosed between 34 and 36 weeks with co-occurring IUGR clustered with the EOPE. DNAm profiling may provide an independent tool to refine clinical/pathological diagnoses into subgroups with more uniform pathology. Despite large changes observed in EOPE, there were challenges in reproducing genome-wide DNAm hits that are discussed.

## Introduction

Preeclampsia (PE) (OMIM 189800), a multi-system maternal hypertensive disorder of pregnancy, is the leading cause of maternal and perinatal morbidity and mortality worldwide, occurring in 2–8% of pregnancies (1). PE is also a major cause of

intrauterine growth restriction (IUGR), defined as poor fetal growth due to an underlying pathology. However, IUGR may also occur in normotensive (non-hypertensive) pregnancies. Infants from pregnancies complicated by PE and/or IUGR are at

Received: September 18, 2017. Revised: September 18, 2017. Accepted: October 26, 2017

© The Author 2017. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

risk for immediate and long-term adverse health outcomes (2,3). To date, there is no consistent test utilized to predict PE or IUGR prior to the onset of clinical symptoms. Protein biomarkers such as pregnancy-associated plasma protein A (PAPPA) and placental growth factor (PLGF) have been used to predict PE and/or IUGR (4); however, these methods may not be generalizable to other populations, as studies consist predominantly of high risk and/or Caucasian populations (5). Another limitation of current screening approaches is our poor understanding of the distinct pathological mechanisms and corresponding placental changes that may underlie these conditions.

Both PE and IUGR are heterogeneous in etiology, with many different factors contributing to these phenotypes (6–8). Risk factors for PE include genetic abnormalities, such as triploidy, trisomy 13 or 16, and point mutations, as well as maternal health factors, such as obesity, pre-existing hypertension, and diabetes (9–12). Normotensive IUGR (nIUGR, i.e. no co-occurring PE) can arise due to similar factors as PE as well as poor maternal nutrition, smoking, stress, and other causes (12–14). Due to both the heterogeneity within and overlap between the etiology of PE and IUGR, the ability to sub-classify placentas into more homogeneous groups should aid in our understanding of disease pathogenesis and its prediction. Abnormal placental findings associated with IUGR and severe preeclampsia are similar and include: placental infarcts of varying types, fibrin deposition, advanced villous maturation, villitis of unknown etiology etc. (15). By defining ‘placental IUGR’ on the basis of a detailed scoring system for placental pathology, Benton *et al.* showed the most abnormal subset was associated with very low maternal serum PLGF and also had the most severe perinatal and postnatal risks (16). In some cases, PE and nIUGR may represent two facets of a common underlying etiology, while in others the associated placental pathology and molecular changes may be distinct. Enforcing stringent criteria for defining and grouping samples may increase the reproducibility for reported molecular changes. For this study, we subdivide our samples into early-onset PE (EOPE), late-onset PE (LOPE), and nIUGR based on clinical obstetric criteria.

Molecular profiling has the potential to refine these clinically-defined group definitions further by identifying the heterogeneity within and the overlap between EOPE, LOPE, and nIUGR. Placental transcriptome profiling from pregnancies associated with PE and healthy controls provide evidence for multiple subtypes of PE (17). DNA methylation (DNAm) profiling is an alternative or complementary approach to gene expression profiling to identify subgroups of placental phenotypes. DNAm is more stable than mRNA and hence is less subject to changes in sample processing time (18); it may also retain a ‘memory’ of earlier in utero exposures and hence be linked to early effects in the disease process.

We previously showed widespread DNAm alterations in EOPE (19) using the Illumina Infinium HumanMethylation450 Array (450K), measuring >480 000 CpG sites across the genome (20). We also demonstrated that placentas associated with confined placental trisomy 16, a condition that can be associated with PE, show some overlapping changes with chromosomally normal EOPE, as well as a unique set of changes specific to the presence of the trisomy (21). Other groups have similarly found altered DNAm in PE and IUGR, though the differentially methylated sites or ‘hits’ are not entirely consistent between studies (22–29). This inconsistency may be due to i) how sample groups are defined, ii) placental sampling differences, or iii) how validated hits are defined between the studies.

The aim of the present study was to evaluate DNAm changes in the context of placental insufficiency due to PE and IUGR as compared with other preterm and term placentas. Specifically, we i) use an epigenome-wide association approach to identify DNAm changes associated with EOPE, LOPE and nIUGR ii) validate differentially methylated sites in an independent cohort and iii) investigate the relationships between PE, IUGR, preterm birth and term control placentas using hierarchical clustering based on the validated hits. We also discuss challenges to validation and the relevance for other applications of epigenetics in the placental biology field.

## Results

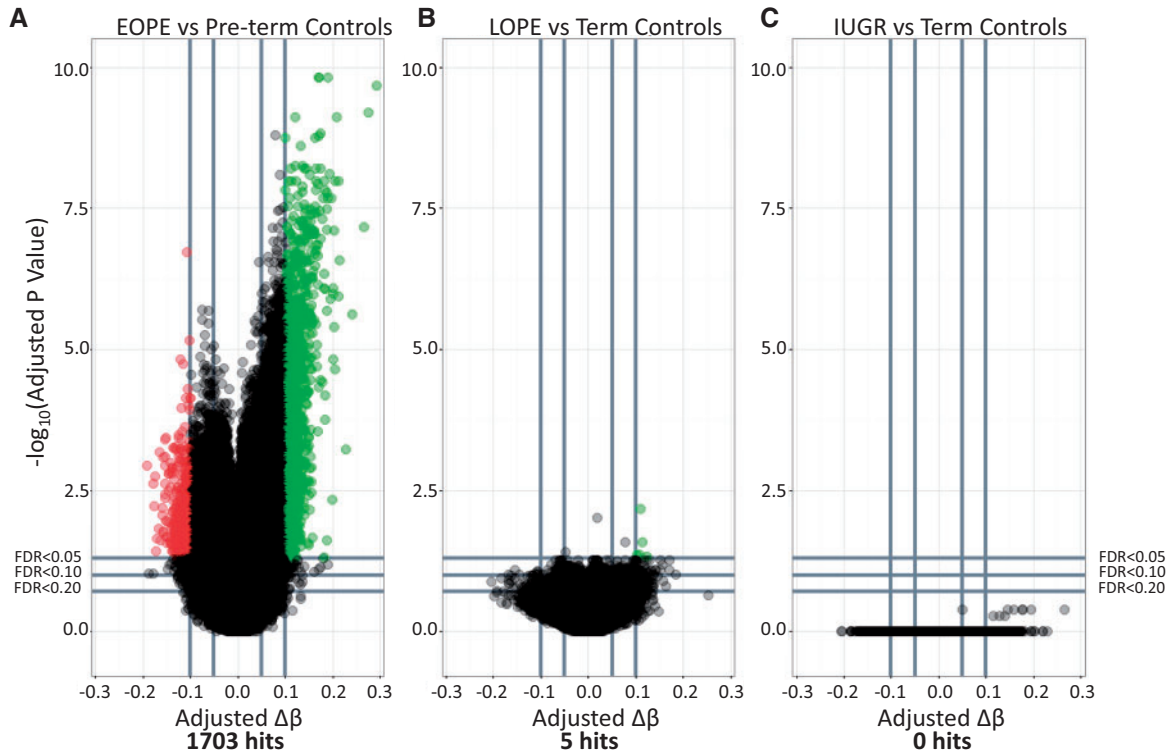
### Widespread DNAm changes are associated with EOPE but not LOPE and nIUGR in our Discovery cohort

Our first goal was to confirm our previous report of widespread changes in EOPE (19) and then to test for similar changes in LOPE and nIUGR using the same approach. We chose less stringent cutoffs for significance in this analysis ( $FDR < 0.05$  &  $\Delta\beta > 0.1$ ) as compared with Blair *et al.* (2013) ( $FDR < 0.01$  &  $\Delta\beta > 0.125$ ) (19) as our aim was to identify a larger number of differentially methylated sites that could be used for further validation. Based on these criteria, a total of 1703 sites were differentially methylated between EOPE and pre-term controls (Fig. 1). As expected, the majority (261/286) of EOPE hits reported in Blair *et al.* were also identified as hits in this analysis. Differences between the two analyses are likely explained by the use of different normalization methods, correction for fetal sex in the present study, and the inclusion of a few additional samples in this study compared with the previous one.

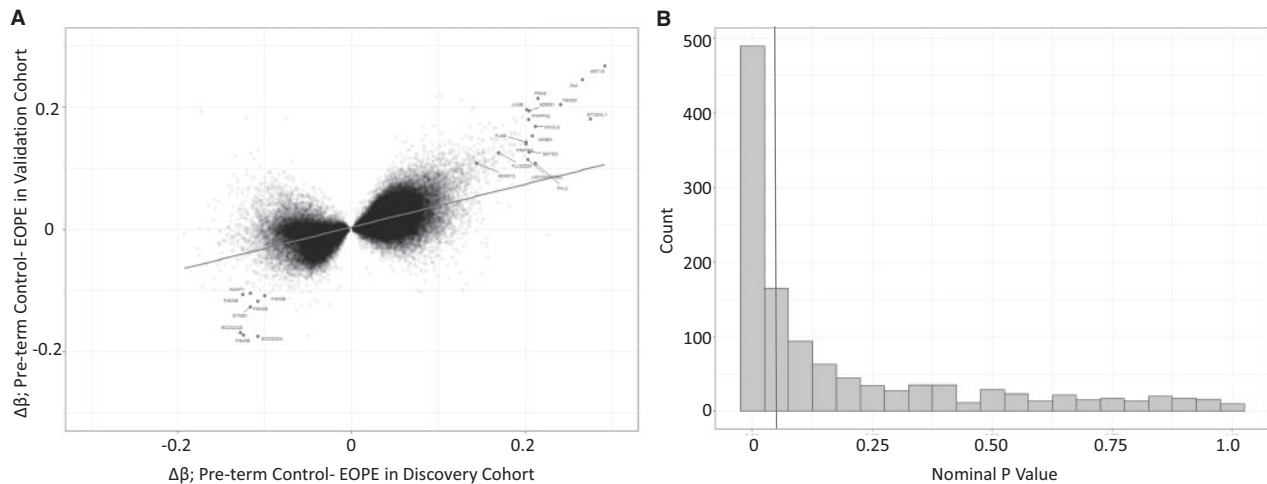
We used the same approach to identify differential methylation associated with LOPE or nIUGR as compared with the healthy term control group. In contrast to the EOPE comparison, only 5 sites were differentially methylated between LOPE and term controls, and no sites were differentially methylated between nIUGR and term controls (Fig. 1). The five differentially methylated sites between LOPE and term controls were not unique to LOPE, as they were also included amongst the 1703 sites identified as differentially methylated in EOPE. These few differentially methylated sites in LOPE may therefore be explained if a subset of the LOPE cases present with typical EOPE pathology.

### Validation of the EOPE hits in an independent cohort

We next investigated if the EOPE hits from our discovery cohort could be validated in an independent cohort. We first tested whether the  $\Delta\beta$  values in the discovery and validation cohorts were correlated using all sites that met an  $FDR < 0.05$  in the discovery cohort, without imposing an additional  $\Delta\beta$  threshold on the validation cohort. At these sites, the correlation was significant ( $R = 0.62$ ,  $P < 2.2e-16$ , Fig. 2A). This indicates that largely similar changes in DNAm are being observed in the EOPE placentas in both cohorts. However, it should be noted that the correlation appeared to be much stronger for sites that had shown a high  $\Delta\beta$  in the discovery cohort, while reproducibility is much poorer for sites that showed small methylation changes. Amongst the most highly hypomethylated sites (Preterm controls- EOPE  $\Delta\beta > 0.15$ ,  $N = 23$ ) in both cohorts were CpGs associated with many of the genes that were previously observed as differentially expressed in preeclampsia *KRT15*, *FN1*, *TEAD3*, *JUNB*, *ST3GAL1*, *PKM2*, *NDRG1*, *PAPPA2*, *CHI3L2*, and *INHBA*.



**Figure 1.** Volcano plots depicting differentially methylated sites between (A) early onset PE (EOPE) and pre-term controls (Pre-term - EOPE), (B) LOPE and term controls (Term - LOPE), and (C) IUGR and term controls (Term - IUGR).  $-\log_{10}$  of the adjusted  $P$ -value is plotted on the y axis and the change in DNAm ( $\Delta\beta$ ) is plotted on the x axis. Sites highlighted in red are hypermethylated in the pathology compared with controls. Sites highlighted in green are those that are hypomethylated in the pathology compared with controls.



**Figure 2.** (A) The correlation between the change in DNAm ( $\Delta\beta$  values) between early-onset PE (EOPE) and pre-term controls, between the discovery and validation cohorts. The sites highlighted are the top sites labeled by the gene associated with the CpG site. (B)  $P$ -value distribution of the 1703 EOPE hits from the discovery cohort, in the validation cohort.

Amongst the most highly hypermethylated sites (Pre-term control- EOPE  $\Delta\beta < -0.10$ ,  $N = 502$ ) in both cohorts were several sites associated with *FAM3B*, *SYNE1*, and *AGAP1*.

To narrow down the original 1703 EOPE hits from the discovery cohort to a high-confidence hit list, we first asked, how many of these hits met similarly stringent criteria ( $FDR < 0.05$  and  $\Delta\beta > 0.1$ ) in the validation cohort? We found that only 38 probes (2.2%) met these strict criteria.

Using such arbitrary cutoffs in both populations and a strict definition for a 'hit' may not be a powerful approach to assess

the degree of overlap in the data. Furthermore, requiring assay-wide correction for multiple testing in the validation cohort is overly conservative and reduces power. Running the linear regression on only the 1703 sites differentially methylated in the discovery cohort reduces the number of multiple test corrections needed in the validation cohort. Based on the distribution of nominal  $P$ -values among the 1703 EOPE associated sites in the validation cohort, shown in **Figure 2B**, there are many more sites that meet a nominal  $P$ -value  $< 0.05$  than expected by chance, even if these do not meet a multiple test correction.

Hence, we opted to use a nominal  $P$ -value  $< 0.05$  and a change in DNAm in the same direction as the discovery cohort to define validated (i.e. high confidence) hits. Based on these criteria, 599 of the 1703 (35.1%) EOPE hits were considered to be validated (Fig. 2B). This is higher than what we would expect by chance ( $P = 0.0001$ ). This reproducibility rate was similar to the rate reported by Yeung et al. (2016), who validated their own differentially methylated regions with our published cohort [Blair et al. (2013)] (30). As PE and IUGR are heterogeneous conditions, it is possible that the reproducibility rate may be affected by the samples chosen for each cohort, as well as sample size. We were interested in whether the samples in the cohorts were similarly correlated (i.e. is one cohort more heterogeneous than the other). We investigated these correlations in the control samples (Term and Pre-term) (Supplementary Material, Fig. S1) and the EOPE samples (Supplementary Material, Fig. S2). In both pathology groups, samples in the discovery cohort were more heterogeneous than the samples in the validation cohort. This heterogeneity suggests that the pre-term birth control placentas used may be different between the discovery and validation cohorts and may explain some of the non-reproduced hits between the two cohorts. This heterogeneity may be due to the fact that the discovery PTB controls were purposely selected to span a range of etiologies and GAs while some pathologies, such as chorioamnionitis, were excluded as controls in the validation cohort.

These validated sites were not enriched for any gene ontology terms using ermineJ, with a 450K array specific background (31). Interestingly, 224/599 validated differentially methylated sites are found in enhancer regions. These sites include ones associated with genes known to be relevant to EOPE from gene expression studies including CGA, INHBA, PAPP2, and ADAM12. A list of these sites and relevant gene information can be found in Supplementary Material, Table S1.

### Effects of varying the validation criteria

To evaluate the effect of varying FDR and  $\Delta\beta$  cutoffs to establish the most 'reproducible' results, we plotted the percentage of probes that showed  $\Delta\beta$  concordance in directionality between the validation cohort and the discovery cohort using different FDRs and  $\Delta\beta$  thresholds in the discovery set (Supplementary Material, Fig. S3A). Different FDR thresholds did not influence DNAm concordance rate when the  $\Delta\beta$  thresholds were above 0.2. FDR thresholds appear to be more important when trying to identify small changes in DNAm. We also investigated the number of hits that each threshold would obtain. Supplementary Material, Figure 3B plots the number of hits at each FDR and  $\Delta\beta$  cutoff. Allowing smaller changes in DNAm produces many more hits, but with a lower reproducibility rate. This is likely because this is in the range of normal variability for a site. This highlights the importance of considering both the biological and statistical thresholds depending on the magnitude of the anticipated DNAm change and the overall research objective.

### Hierarchical clustering

Next, we evaluated the degree to which the 599 validated sites can discriminate EOPE from all other placentas in both cohorts, using hierarchical clustering (Fig. 3). We applied hierarchical clustering to both the discovery and validation cohorts on the 599 validated sites. Interestingly, both cohorts clustered into three stable methylation clusters (determined by 1000

permutations with *pvclust* and *sigclust2*; see methods), rather than just two as we had expected (Fig. 3). Methylation cluster 1 included almost all EOPE, suggesting a consistent 'placental mediated' phenotype in this group. In the discovery cohort cluster, 6 of the 18 LOPE samples clustered with the EOPE samples. In the validation cohort, 7 of the 11 LOPE samples and 1 pre-term control clustered with the EOPE samples. Additionally, in both cohorts, stable and significantly different sub-clusters were identified within the larger placental mediated phenotype group (methylation cluster 1), suggesting that methylation cluster 1 may be heterogeneous itself. However, we observed no obvious difference between these subclusters clinically (including sex, ethnicity, disease severity etc.). Gestational age was decreased in the EOPE methylation subcluster 1 compared with subcluster 2 ( $P < 0.01$ ) in the validation cohort (Table 1). As we could not identify a clinical reason why the subclusters existed and that sample size in the subclusters was small within methylation cluster 1, we did not split them into separate groups for further analysis.

The remaining non-EOPE samples also separated into two methylation clusters in both cohorts. We refer to these clusters as methylation clusters 2 and 3. Methylation cluster 3 in both cohorts was predominantly composed of controls. Within the discovery cohort, methylation cluster 2 consisted of the majority of the nIUGR and LOPE cases, a few EOPE cases, and some pre-term and term controls. Additionally, decreased birthweight ( $P < 0.01$ ) and a trend towards increased IUGR diagnosis ( $P < 0.1$ ) was observed in methylation cluster 2 compared with cluster 3. In the validation cohort, methylation cluster 2 consisted of pre-term controls and LOPE samples and was associated with decreased gestational age ( $P < 0.01$ ) (Table 1).

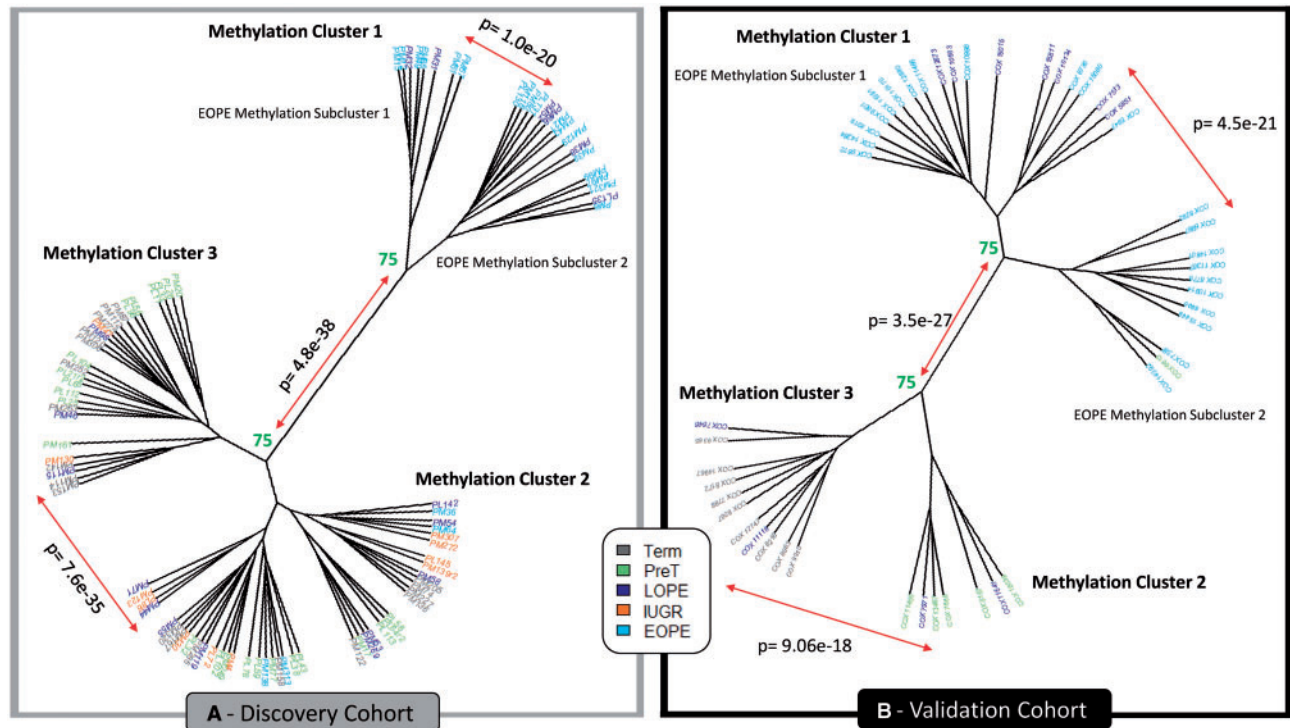
### Cluster gene ontology

As it was unexpected that the control samples would split into two distinct methylation clusters, we were interested in investigating the differences between the two control methylation clusters (methylation clusters 2 and 3). Between methylation cluster 2 and methylation cluster 3, 244 sites were differentially methylated in both cohorts. Information on these sites can be found in Supplementary Material, Table S2. There was no gene ontology enrichment by ermine or DAVID. Between EOPE methylation subcluster 1 and EOPE methylation subcluster 2, 207 sites were differentially methylated in both cohorts. Information on these sites can be found in Supplementary Material, Table S3. There was no gene ontology enrichment by ermineJ and *symporter* activity was the only gene ontology term in DAVID to meet multiple test corrections.

### Discussion

We previously reported widespread changes in DNAm associated with EOPE (19). In the present study, we extend this analysis to LOPE and nIUGR; however, using the same approach, we were unable to identify DNAm changes that are unique to these groups. While these latter comparisons were limited by small sample size, we were able to obtain significant associations with EOPE with similarly small sample sizes. The reduced number of changes in the LOPE and nIUGR groups can occur for two main reasons: 1) there may be much more limited placental pathology with these diagnoses and the phenotype is largely driven by maternal factors, or 2) they may be more heterogeneous etiology thereby limiting power to detect changes in the





**Figure 3.** Hierarchical clustering (Euclidean) on the 599 validated hits in both the discovery (left) and validation (right) cohorts. Numbers represent the percentage of times these clusters formed when using 1000 iterations with pvclust. Those highlighted in green are considered stable, where clusters formed >75% of the time. Those highlighted in red were unstable. P values signify clusters are significantly different from one another.

**Table 1.** Clinical information on samples assigned to methylation cluster 2 compared with methylation cluster 3 and samples assigned to EOPE methylation subcluster 1 and EOPE methylation subcluster 2. \*p-value<0.1, \*\*p-value<0.05, \*\*\*p-value<0.01.

#### DISCOVERY COHORT

|   | Methylation Cluster 2 | Methylation Cluster 3 | EOPE Methylation Subcluster 1 | EOPE Methylation Subcluster 2 |
|---|-----------------------|-----------------------|-------------------------------|-------------------------------|
| N =   | 43                    | 27                    | 8                             | 16                            |
| IUGR Status (IUGR/Total)                                | 10/43 (23%)*          | 2/27 (7%)             | 7/8 (88%)                     | 14/16 (88%)                   |
| Fetal Sex (F/Total)                                     | 22/43 (51%)           | 11/27 (41%)           | 5/8 (63%)                     | 3/16 (19%)*                   |
| Gestational Age (weeks) [range (mean)]                  | 25.0–40.0 (36.0)      | 28.0–41.3 (35.7)      | 24.9–36.0 (31.7)              | 26.0–37.3 (32.6)              |
| Fetal Birth Weight (SD) [range (mean)]                  | –2.78–3.77 (–0.44)**  | –2.16–1.70 (0.18)     | –2.90–1.17 (–1.87)            | –8.19 – -0.58 (–2.26)         |
| Chronic Hypertension (CH/Total)                         | 7/43 (16%)            | 2/27 (7%)             | 2/8 (25%)                     | 5/16 (31%)                    |
| Diabetes (Pre-existing or Gestational) (Diabetes/Total) | 2/43 (5%)             | 0/27 (0%)             | 1/8 (13%)                     | 1/16 (6%)                     |
| Chorioamnionitis (CA/Total)                             | 7/43 (16%)            | 5/27 (19%)            | 0/8 (0%)                      | 0/16 (0%)                     |
| Premature Rupture of Membranes (PPROM/Total)            | 3/43 (7%)             | 4/27 (15%)            | 0/8 (0%)                      | 0/16 (0%)                     |
| Ultrasound Findings (Findings/Total)                    | 7/43 (16%)            | 2/27 (7%)             | 3/8 (38%)                     | 4/16 (25%)                    |
| Placental Pathology Noted (Notes/Total)                 | 17/43 (40%)           | 7/27 (26%)            | 4/8 (50%)                     | 10/16 (63%)                   |

#### VALIDATION COHORT

|  | Methylation Cluster 2 | Methylation Cluster 3 | EOPE Methylation Subcluster 1 | EOPE Methylation Subcluster 2 |
|--|-----------------------|-----------------------|-------------------------------|-------------------------------|
| N =                                    | 7                     | 11                    | 19                            | 11                            |
| IUGR Status (IUGR/Total)               | 0/7 (0%)              | 0/11 (0%)             | 11/19 (58%)                   | 7/11 (64%)                    |
| Fetal Sex (F/Total)                    | 1/7 (14%)             | 5/11 (45%)            | 9/19 (47%)                    | 7/11 (64%)                    |
| Gestational Age(weeks) [range (mean)]  | 30.0–37.0 (33.0)***   | 37.0–40.0 (38.4)      | 27.0–37.0 (32.8)***           | 26.0–34.0 (29.8)              |
| Fetal Birth Weight (SD) [range (mean)] | –0.48–0.50 (–0.06)    | –1.11–3.60 (0.34)     | –2.27–2.92 (–1.12)            | –2.46 – -0.86 (–1.59)         |
| Chronic Hypertension (CH/Total)        | 1/7 (14%)             | 1/11 (9%)             | 7/19 (37%)                    | 3/11 (27%)                    |

group as a whole. In fact, 6 of the 18 LOPE samples in the discovery cohort and 7 of 11 in the replication cohort were later shown by a clustering approach to have DNAm profiles similar to the EOPE group, which illustrates the importance of having homogeneous pathology groups for detecting pathology-associated DNAm changes. If we want to improve biomarker discovery in these groups, we will need to identify more homogeneous subgroups using a combination of clinical parameters, pathology reports and/or biomarkers themselves, along with larger sample sizes.

The LOPE samples in the discovery cohort that clustered with the EOPE samples all presented with PE between 34.0 weeks and 35.9 weeks gestation and had co-occurring IUGR. While it is possible that PE symptoms were present but not diagnosed until after 34.0 weeks, there may also be inaccuracies in dating the pregnancy and/or there is simply a grey zone in the distinction between EOPE (placenta-driven) and LOPE (maternal-health driven). There were also four cases of EOPE within the discovery cohort that did not cluster with other EOPE cases. One was diagnosed with hemolysis elevated liver enzymes and low platelet (HELLP) syndrome and delivered at 33.3 weeks gestation, one also had chorioamnionitis (which may have contributed to early delivery); one had preexisting hypertension and was diagnosed early but did not deliver until 37 weeks and hence may have been milder in presentation; the fourth was delivered at 33.3 weeks with no other placenta or maternal health notes. Of note, none of the EOPE cases that clustered outside of the EOPE cluster had co-occurring IUGR and they were generally diagnosed at close to the cutoff of 34 weeks gestation. Thus, the presence of IUGR in PE cases specifically diagnosed between 32 and 36 weeks may be the more defining feature as to whether an altered placental DNAm profile is observed or not. Our data suggest that PE that co-occurs with IUGR is a distinct entity from both isolated PE (no IUGR) and isolated IUGR (no PE). Powers *et al.* (2012) showed that there are two types of PE pregnancies: those with and without altered angiogenic factors (32). As alterations in angiogenic factors have also been observed in IUGR cases (33,34), Myatt and Roberts suggested that an imbalance in these factors may represent a measure of placenta growth, development, and function (35). As such, the EOPE cases clustering outside the EOPE methylation cluster may be more likely related to other contributing factors than placental dysfunction.

While we expected to see an EOPE methylation cluster using the validated hits chosen based on being differentially methylated between EOPE and pre-term controls, we were surprised that both the EOPE and control groups each formed two subclusters. The driving differences between these subclusters were not clear, though the tendency to lower gestational ages and fetal birth weights (Standard deviation; SD) in subcluster 2 could suggest features linked to preterm birth (Table 1). When looking specifically at sites that were differentially methylated between methylation cluster 2 and methylation cluster 3, and between methylation cluster 1 and methylation cluster 3, we noticed some genes with similar functions, some of which have previously been associated with PE. *FN1*, *PKM2*, and *KRT15* had changes in DNAm  $>0.1$  and FDR  $<0.05$  in both the validation and discovery cohorts between methylation cluster 1 and methylation cluster 3. *FN1* is involved in cell adhesion (36) and protein levels in maternal blood have been reported increased in EOPE and LOPE with IUGR pregnancies, but not in LOPE without IUGR or nIUGR (37). *PKM2* is involved in cellular metabolism, gene expression levels on *PKM2* have been reported increased in placentas of PE and IUGR pregnancies (38). *KRT15* encodes

keratin important to cellular structure and integrity (39). *KRT15* has been reported absent from cytotrophoblast vesicles in pregnancies complicated by PE (40). Figure 4A–C show that methylation cluster 1, likely consisting of ‘placental mediated PE’ is hypomethylated at sites within *FN1*, *PKM2*, and *KRT15*. DNAm distributions at these sites show a clear division between methylation cluster 1 and the other clusters. *IL7*, *MAF*, and *CXCL9* were differentially methylated between methylation cluster 2 and methylation cluster 3 in both cohort (FDR  $<0.05$ ,  $\Delta\beta >0.05$ ) and all have immune related functions (41–43). The beta value distribution of methylation cluster 2 appeared to have an intermediate DNAm distribution between methylation cluster 1 and 3 for all of these 6 sites (Fig. 4). Clustering on just these 6 CpG sites, principal component analysis clearly distinguished our 3 methylation clusters (Supplementary Material, Fig. S5). The presence of two distinct subclusters within the EOPE methylation cluster could reflect PE severity, or perhaps unmeasured factors such as medical treatments given or duration of hypertension. Unfortunately, we had insufficient information on the treatment of each case to evaluate the influence of medical care on the placental methylation profile. The two clusters would also be a technical artifact as our current information cannot identify a biological reason for the split in methylation cluster 1.

Our data show a correlation in the changes in DNAm between EOPE cases in the discovery cohort compared with control cohort (Fig. 2A). However, it should be noted that there were also many sites with a high  $\Delta\beta$  in the discovery cohort that had a much smaller or sometimes opposite direction  $\Delta\beta$  in the validation cohort. These sites were not significantly enriched for any genomic region, but lack CpG sites in enhancer regions (Supplementary Material, Fig. S4). Lack of validation could be due to i) differences between the cohorts; ii) violating statistical assumptions, or iii) technical noise. For example, risk factors for PE include maternal health, lifestyle factors, genetic ancestry etc., which could differ between the Vancouver and Toronto cohorts and influence DNAm. Additionally, when performing an FDR analysis, we assume the distributions of DNAm for each probe is similar and independent. In fact, many sites show correlated changes and so are not independent and some sites are hypervariable and influenced by genetic factors that could differ between populations. Our results also show that reproducibility is greater in sites where the change in DNAm higher, with changes in DNAm  $>10\%$  points being more robust. This emphasizes the need for validation of changes as there are multiple unknown reasons why a strong result in one cohort may not replicate in another.

While altered placental DNAm has been reported for pregnancies complicated by PE (22–24,26) and IUGR (25,27,28), only one study validated their findings using the same technology in an independent cohort (30). In this study, 35.1% ( $N=599$ ) sites were found to be differentially methylated between EOPE and pre-term controls in both the discovery and validation cohorts using validation criteria of a nominal  $P < 0.05$  and change in DNAm in the same direction as the discovery cohort. The extent of validation, however, is dependent on the initial criteria chosen to define ‘hits’, the criteria for validation, the similarity of the populations of samples, the size of the study populations (power to detect changes), and the similarity in processing the samples. In our study, the validation cohort was from a roughly similar urban population (Vancouver vs. Toronto) from the same country (Canada) (44). Cases with self-reported ancestry in the discovery cohort (~35% of data) consisted of ~60% Caucasian, 30% Asian and 10% other. While this represents to some extent the Vancouver population, there is likely an

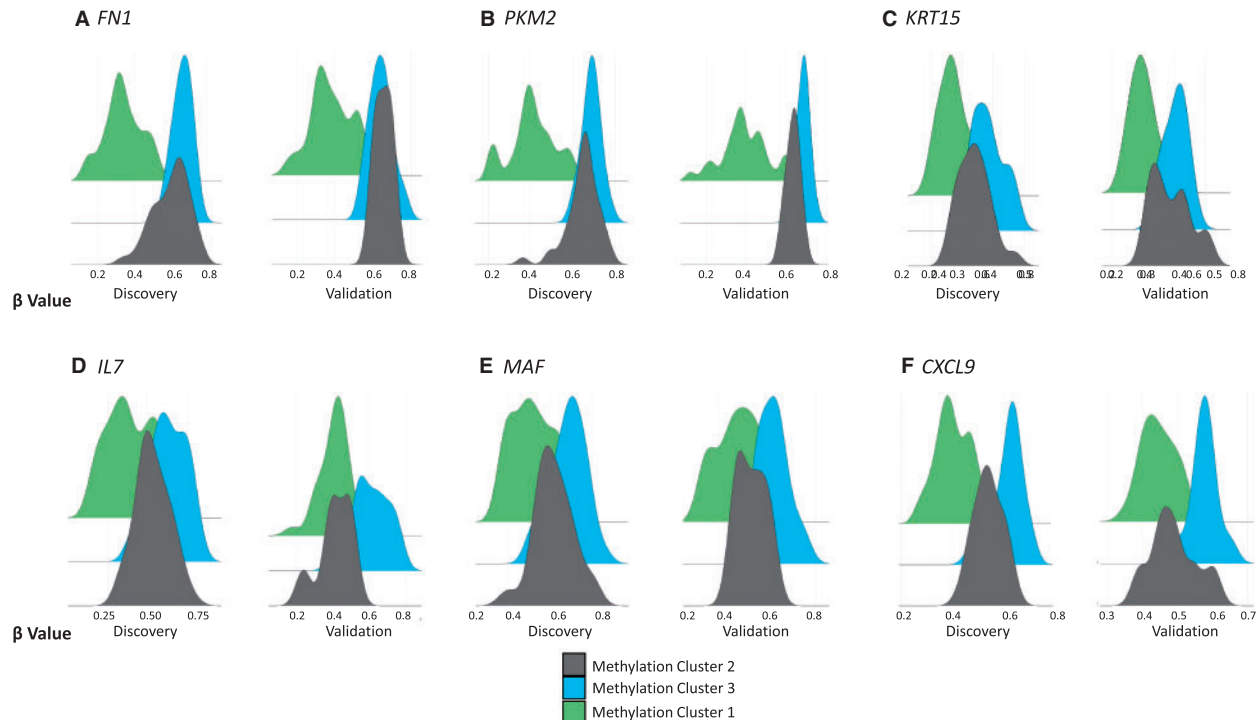


Figure 4. Beta value distributions by methylation cluster for (A) *FN1*, (B) *PKM2*, (C) *KRT15*, (D) *IL7*, (E) *MAF*, and (F) *CXCL9*.

overrepresentation of Caucasians as they are more likely to consent to study participation. From the self-reported ethnicity in the validation cohort, 52% were Caucasian, 25% Asian, 15% Black, and 8% other. We also tried to minimize technical factors that may influence results by using similar placental sampling protocols, processing the arrays with a subset of the discovery and validation cohorts on the same microarray chips at the same time, with the same technicians, and using the same preprocessing methods on the raw data. Even with these considerations, a significant number of our original hits were not validated. This may be because of chance variation in causes of PE and IUGR in the two cohorts due to limited sample size. Additionally, there are genetic, environmental, and maternal factors that pre-dispose a pregnancy to developing placental insufficiency, which may have varied between populations.

Changes in DNAm could mean i) an average change in DNAm across a sample, or ii) a change in the cell type proportions within a sample, as DNAm varies widely across different cell types (45). In the context of EOPE, DNAm alteration may reflect a combination of altered gene expression pathways associated with PE [ex. related to known effects such as oxidative stress and altered angiogenesis (7,32)] or altered cell type proportions related to PE pathology [ex. decreased proliferation of extravillous trophoblast cells or alterations to the rate of trophoblast proliferation (46)]. Therefore, correcting for cell composition in this context (PE and IUGR) may remove variance due to pathology as pathology and cell composition may be confounded. As cell type-specific profiles have not been developed for all placental cell types, it is not possible to use the DNAm profile to estimate cell proportions, as it has been applied to blood (47). While reference-free methods for deconvolution of cell proportions have been developed (48), these methods remove variance within the data attributed to cell composition but cannot inform us of what cell types specifically are altered in EOPE.

## Summary

Our data demonstrate some of the challenges in identifying changes specific to clinically defined etiologies. Heterogeneity and milder phenotypes of LOPE and nIUGR likely limit the power to detect differences using a differential methylation type approach and mask the subset of cases that do exhibit altered pathology (based on sample clustering using our EOPE defined hits). An alternative approach may be to reduce the dimensions in the data by i) removing non-variable probes across all cell types (49), ii) focusing on alterations in pathway modules, as in weighted gene co-expression network analysis (50), or iii) evaluating differentially methylated regions (DMRs) rather than individual CpG sites to combat multiple test correction (51,52). In contrast, the more severe pathology underlying EOPE results in many readily detected DNAm changes. However, even in the case of EOPE, where many large changes in DNAm are identified and can be validated based on a nominal  $P$ -value  $< 0.05$  in an independent cohort, those sites selected for having the highest magnitude of change rarely showed the same degree of difference in the second cohort. Techniques to reduce the dimensions in the data should be developed and utilized, focusing on altered pathways instead of specific changes, which may help in identifying subtypes of PE and IUGR to guide and change management in a useful way.

In conclusion, whether in the context of PE, or other heterogeneous diseases, DNAm may be a useful tool to independently and qualitatively classify pathological groups, after delivery. This subclassification into more homogenous groups will assist in better prediction of post-natal effects of *in utero* conditions, and is a necessary precursor to develop predictive screening tools specific to each phenotype to identify high-risk pregnancies (53). The developmental origins of health and disease (DoHaD) seeks to understand how *in utero* exposures affect long term health of the infant, and many have associated IUGR with

increased risk of adult onset diseases (54,55). Understanding the cause of these more homogenous groups of PE and IUGR is imperative to clinical management of subsequent pregnancies and infants who may have an increased risk for adult onset diseases.

## Materials and Methods

### Sample information

#### Discovery cohort

Chorionic villi samples were obtained from 2 to 3 sites in the placenta, each from distinct cotyledons, as previously described (19). Infarcts or necrotic regions of the placenta were avoided in sampling and DNA was extracted from each site, using the Qiagen blood and tissue kit and pooled together in equal amounts to give a more accurate representation of the placenta's molecular profile.

The discovery (Vancouver) cohort consisted of 22 EOPE, 18 LOPE, 11 nIUGR and 43 control placentas (Table 2). Ethics approval from both the University of British Columbia and BC Women's and Children's Hospital ethics committees in Vancouver, BC, Canada, was obtained (H04-704488). Placental samples were obtained with consent from patients from the Medical Genetics as well as the Obstetrics and Gynecology departments. Clinical information, including gestational age at delivery, fetal sex, fetal birth weight, and maternal age were collected. Criteria for exclusion were multi-fetal pregnancies and fetal and/or placental chromosomal abnormalities. A subset of 18 EOPE samples and 19 pre-term control samples in this study was previously used in Blair *et al.* (2013) (19).

PE was defined according to the Society of Obstetricians and Gynecologists of Canada (SOGC) criteria as one of i) hypertension (BP > 140/90mm Hg) and proteinuria (>300mg/day) arising after 20 weeks gestation (56); ii) HELLP syndrome without hypertension or proteinuria; or iii) eclamptic seizure without previous hypertension or proteinuria. Preeclampsia was separated into early and late onset given the clinical evidence that these may be associated with distinct risk factors and outcomes (6). Early-

onset preeclampsia (EOPE) was defined as a diagnosis of PE prior to 34 weeks gestation, while LOPE was PE diagnosed after 34 weeks (6). nIUGR was defined as fetal birth weight <3<sup>rd</sup> percentile accounting for both fetal sex and gestational age at delivery or fetal birth weight <10<sup>th</sup> percentile accounting for both fetal sex and gestational age at delivery, with additional findings for poor fetal growth (56) and no hypertension. As birthweight is strongly correlated with gestational age, we use the standard deviation of the birth weight corrected for fetal sex and gestational age (57). In majority of cases, gestational age was dated using ultrasound measures which is standard practice in Canada.

Technical batch effects, related to the plate, microarray chip, and sample position on the Illumina chip are potential confounding factors within our data. Our samples were run in various batches over a 4 year period, and pathology and gestational age were partially confounded with batch as EOPE and preterm controls were largely run earlier. In this situation, correction for batch effects can introduce spurious findings (58) (Supplementary Material, Figs S6 and S7). We, therefore, instead compared EOPE to pre-term birth controls and LOPE/nIUGR to term controls only, which were relatively matched for batch, and thus the confounding by GA and its interaction with batch was minimized. We acknowledge that some of the differentially methylated sites that we found may be due to technical artifacts, but focusing on those hits that are reproduced in the validation cohort largely eliminated these effects.

As placental DNAm changes with gestational age, the comparison groups included placentas from healthy term births ( $\geq 37$  weeks) and pre-term births (<37weeks) with normally grown babies and no evidence of maternal hypertension. EOPE placentas were compared with 24 pre-term controls (as in Blair *et al.* 2013). LOPE and nIUGR placentas were compared with a separate set of 19 term control placentas. This was to test for overlap between DNAm changes identified for LOPE and nIUGR with those for EOPE, as we did not want the use of a control group driving any potential overlap. To reduce the chance of differences being driven by the preterm birth group, we used

**Table 2.** Discovery and validation cohort clinical information

| Discovery Cohort                             |                                 |                                 |                                 |                                 |                     |
|--|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------|
|  | EOPE (mean)                     | LOPE (mean)                     | IUGR (mean)                     | Pre-term 'Control' (mean)       | Term Control (mean) |
| N=   | 22                              | 18                              | 11                              | 24                              | 19                  |
| Fetal birth weight- Standard deviation range | -8.18-3.77 (-1.65) <sup>1</sup> | -2.9-2.57 (-0.96) <sup>2</sup>  | -2.57-1.22 (-1.99) <sup>3</sup> | -1.61-3.23 (0.51)               | -0.94-0.98 (-0.09)  |
| Fetal Sex (M: F)                             | 14: 8                           | 10: 8                           | 4: 7                            | 16: 8                           | 9: 10               |
| Maternal Age (years)                         | 19.7-42.9 (33.3)                | 23.1-41.3 (34.0)                | 33.3-38.0 (34.3)                | 22.2-41.1 (32.5)                | 30.0-40.2 (34.9)    |
| Gestational Age (weeks)                      | 24.9-38.4 (32.0)                | 34.6-41.3 (37.4)                | 34.6-38.0 (36.6) <sup>4</sup>   | 25.0-36.7 (32.6)                | 37.3-39.9 (38.4)    |
| Validation Cohort                            |                                 |                                 |                                 |                                 |                     |
|  | EOPE (mean)                     | LOPE (mean)                     | IUGR (mean)                     | Pre-term 'Control' (mean)       | Term Control (mean) |
| N=   | 22                              | 11                              | 0                               | 6                               | 9                   |
| Fetal birth weight- Standard deviation range | -2.46-0.01 (-1.40) <sup>5</sup> | -2.92-0.01 (-0.84) <sup>6</sup> | 0                               | -1.02-0.05 (-0.16) <sup>7</sup> | -1.11-3.60 (0.61)   |
| Fetal Sex (M: F)                             | 9: 13                           | 7: 4                            | 0                               | 4: 2                            | 6: 3                |
| Gestational Age (weeks)                      | 26.0-34.0 (30.5)                | 35.0-37.0 (36.4) <sup>8</sup>   | 0                               | 27.0-33.0 (30.8)                | 38.0-40.0 (38.8)    |

1) p-value=6.9e-6 vs. pre-term control, 2) p-value=0.014 vs. term control, 3) p-value=1.7e-5 vs. term control, 4) p-value=3.3e-4 vs. term control, 5) p-value=0.004 vs. pre-term control, 6) p-value=0.0005 vs. term control, 7) p-value=0.04 between discovery cohort pre-term control vs. validation cohort pre-term control, and 8) p-value=0.0001 vs. term control.



placentas from pre-term births from a variety of etiologies (e.g. Premature rupture of the membranes, incompetent cervix, chorioamnionitis), while any term control samples with evidence of pathology involving the chorionic villi were excluded. There was a significant difference in gestational age between nIUGR and term placentas, as well as between nIUGR and pre-term placentas. We opted to compare nIUGR to term controls, as the gestational ages were closer in range to the term controls, acknowledging the changes in DNAm identified may be due to gestational age. Gestational age was to be assessed on identified sites *post hoc*, however, as no differentially methylated sites were identified between nIUGR and controls, this analysis was not completed and focus was shifted to validated EOPE differentially methylated sites.

#### Validation cohort

The validation (Toronto) cohort consisted of 22 EOPE, 11 LOPE, and 15 control placentas (Table 2). For the validation cohort, placental samples were purchased through the Research Centre for Women's and Infants' Health BioBank (Mount Sinai Hospital); details in the sample processing can be found in Leavey *et al.* (2016) (59). DNA was extracted from the pooled placental tissue by ethanol precipitation using the Wizzard<sup>®</sup> Genomic DNA Purification Kit (Promega). Gestational age at delivery and fetal birth weight were collected for each case. Gestational age was dated primarily using second trimester ultrasound markers. For this cohort, ethics approval was obtained from both Mount Sinai Hospital (#13-0211-E) and the University of Toronto (#29435).

The validation cohort represents a subset of samples from the Leavey *et al.* (2016) study (59). PE was defined as BP > 140/90mm Hg after 20 weeks gestation and proteinuria >300mg/day or >2+ by dipstick (60), this is the same as the SOGC guidelines used to define PE in the discovery cohort. As the time of diagnosis was unknown, we subdivided the PE samples from this cohort into EOPE and LOPE based on the gestational age at delivery. Exclusion criteria included diabetes, sickle cell anemia, morbid obesity, and multi-fetal pregnancies. The division of the term and pre-term controls was also done in the validation cohort, which consisted of 6 pre-term control and 9 term control placentas (Table 2).

#### DNA methylation analysis

The NanoDrop 1000 spectrophotometer (ThermoScientific, Wilmington, DE, USA) was used to assess DNA purity and concentration, and 750ng of DNA from each placenta was bisulfite converted using the EZ DNA Methylation Kit (Zymo Research, Irvine, USA). Samples were run on the Illumina HumanMethylation450 BeadChip array platform (450K), measuring DNAm at 485, 512 CpG sites across the genome (20). The samples, run on 27 chips in 4 batches, were hybridized to the microarray chip as per the manufacturer's protocol, and microarray chips were scanned by a HiScan 2000 (Illumina). To minimize any effects of sample processing, validation cohort arrays were run in the same batch and with the same operators as a subset of the samples from the discovery cohort (Chips 5013, 5015, 3024, 3037, 3038, 3110, See Supplementary Material, Figs S6 and S7). This DNA methylation data for the discovery and validation cohorts is available from the Gene Expression Omnibus (GEO) database under the accession numbers GSE100197 and GSE98224, respectively.

Raw data (IDAT Files) were read into R statistical software, version 3.2.4, where functional normalization (61), background

subtraction, and colour correction were performed. Blair *et al.* (2013), previously used subset within-array normalization (SWAN). Functional normalization performs all the benefits of SWAN normalization and, in addition, utilizes the 848 control probes on the array to mediate changes in DNAm that are due to technical effects (61). To check that all sample-reported fetal sexes matched the XY chromosomes in the data, samples were clustered on probes intensities from the XY chromosomes. All reported fetal sexes matched the presence of XY chromosomes in both cohorts. Bad quality probes and those that had a missing beta value in > 5% of samples or a detection P-value < 0.01 were removed from the analysis (Discovery N = 1402, Validation N = 1115). To minimize fetal sex effects, probes on the X and Y chromosomes (Discovery N = 11 648, Validation N = 11 302), as well as probes that cross-hybridize to the X and Y chromosomes (Discovery N = 11 412, Validation N = 10 734), and probes containing a SNP at the CpG of interest were also removed (Discovery N = 19 957, Validation N = 20 398) (62). This left 440 093 CpG sites for analysis in the discovery cohort and 441 963 CpG sites in the validation cohort (Supplementary Material, Table S4).

#### Differential methylation analysis

All statistical analyses were performed using R version 3.2.4. Differentially methylated sites were identified using statistical, i.e. false discovery rate (FDR) < 0.05, and biological, i.e. a change in DNAm ( $\Delta\beta$ ) > 0.1, criteria. We corrected for fetal sex in our linear regression (using *limma* package in R), but we did not adjust for fetal birth weight, as it is closely related to pathology. As our groups were matched to controls of a similar gestational age, our final model where DNAm alterations were identified took into account fetal sex only. Only those sites that met both these criteria were then evaluated in the validation cohort. In this case, linear regression (using the *glm* package in R) was used, and sites were considered to be persistent hits if the nominal P-value < 0.05 and the change in DNAm was in the same direction as the discovery cohort. Bonferroni correction  $P < 0.05$  was also used to investigate how many hits would be validated with a more stringent threshold. To compare the heterogeneity of the samples in each cohort, sample-by-sample Pearson correlations were performed, and the average correlation of each sample was compared between the two cohorts by Student's t-test separately for the control (Term + Pre-term) and EOPE samples.

To investigate whether the 42 Bonferroni corrected hits and the 599 nominal P-value validated hits were more than would be expected by chance, 1703 sites (number of EOPE hits in the discovery cohort) were randomly sampled from the validation cohort data and run through a linear model, correcting for fetal sex. One thousand permutations were run and the number of sites that met a nominal P-value < 0.05 in each iteration was recorded. The number of randomly sampled sites to meet a nominal P-value < 0.05 were compared with the actual number of sites that validated in our data [ $N = 42$  (Bonferroni corrected) and  $N = 599$  (nominal P-value)].

#### Clustering analysis

Hierarchical clustering was performed on the persistent hits to investigate whether samples clustered according to their clinically diagnosed pathology, or whether DNAm profiling could suggest an improved definition of pathological groups. The *pvClust* package in R (63) assessed how stable any resulting clusters were, using 1000 iterations. The *sigClust2* package (64) determined if any clusters were significantly different from one another, also using 1000 iterations. To investigate whether

differences in DNAm between the clusters were enriched for any specific pathway(s), linear regression was used to identify differentially methylated sites between clusters. Differentially methylated sites were annotated to genes using the Price et al. annotated closest transcriptional start site (62), and then inputted into ermineJ, a gene ontology tool (31). ErmineJ allows us to input a background gene list specific to the Illumina 450K array, accounts for multifunctionality (gene ontology terms that appear frequently due to the number of genes involved in the pathway), and allows for multiple iterations to be run to strengthen the power of the analysis.

## Supplementary Material

Supplementary Material is available at HMG online.

## Acknowledgements

We would like to thank Kristal Louie, Dr. J Schuetz, Dr. S Langlois, and Dr. P von Dadelszen for recruiting patients for this study; Ruby Jiang for doing the placental dissections and DNA extractions, and Dr. M Penaherrera with help in both patient recruitment, and assistance in running the 450K arrays. Thanks to Dr. Michael Kobor for the use of the HiScan Machine and facilities to run the 450K arrays. Thanks to all members of the Robinson lab for reviewing and providing valuable feedback on the manuscript.

Conflict of Interest statement. None declared.

## Funding

Canadian Institutes of Health Research (CIHR) operating grant to WPR [#49520 to WPR] and to BC [#128369 to BC]. WPR receives salary support through an investigatorship award from the BC Children's Hospital Research Institute (BCCHRI). SLW is funded through the University of British Columbia Four Year Doctoral Fellowship, KL is funded through an Ontario Graduate Scholarship, and BC is funded through a Tier II Canada Research Chair. Funding to pay the Open Access publication charges for this article was provided by Canadian Institutes of Health Research (CIHR).

## References

1. Steegers, E.A., von Dadelszen, P., Duvekot, J.J. and Pijnenborg, R. (2010) Pre-eclampsia. *Lancet*, **376**, 631–644.
2. Suhag, A. and Berghella, V. (2013) Intrauterine growth restriction (IUGR): Etiology and diagnosis. *Curr. Obstet. Gynecol.Rep.*, **2**, 102–111.
3. Lin, S., Leonard, D., Co, M.A., Mukhopadhyay, D., Giri, B., Perger, L., Beeram, M.R., Kuehl, T.J. and Uddin, M.N. (2015) Pre-eclampsia has an adverse impact on maternal and fetal health. *Transl. Res.*, **165**, 449–463.
4. Poon, L.C., Syngelaki, A., Akolekar, R., Lai, J. and Nicolaides, K.H. (2013) Combined screening for preeclampsia and small for gestational age at 11–13 weeks. *Fetal. Diagn. Ther.*, **33**, 16–27.
5. Audibert, F., Boucoiran, I., An, N., Aleksandrov, N., Delvin, E., Bujold, E. and Rey, E. (2010) Screening for preeclampsia using first-trimester serum markers and uterine artery doppler in nulliparous women. *Am. J. Obstet. Gynecol.*, **203**, 383.e1–383.e8.
6. Von Dadelszen, P., Magee, L.A. and Roberts, J.M. (2003) Subclassification of preeclampsia. *Hypertens. Pregnancy*, **22**, 143–148.
7. Roberts, J.M. and Bell, M.J. (2013) If we know so much about preeclampsia, why haven't we cured the disease?. *J. Reprod. Immunol.*, **99**, 1–9.
8. Villar, J., Carroli, G., Wojdyla, D., Abalos, E., Giordano, D., Ba'aqeel, H., Farnot, U., Bergsjø, P., Bakketeig, L., Lumbiganon, P. et al. (2006) Preeclampsia, gestational hypertension and intrauterine growth restriction, related or independent conditions?. *Obstet. Gynecol.*, **194**, 921–931.
9. Bartsch, E., Medcalf, K.E., Park, A.L. and Ray, J.G. and High Risk of Pre-eclampsia Identification Group (2016) Clinical risk factors for pre-eclampsia determined in early pregnancy: Systematic review and meta-analysis of large cohort studies. *BMJ*, **353**, i1753.
10. Boyd, P., Lindenbaum, R. and Redman, C. (1987) Pre-eclampsia and trisomy 13: A possible association. *Lancet*, **2**, 425–427.
11. Yong, P.J., Langlois, S., Dadelszen, P.v. and Robinson, W. (2006) The association between preeclampsia and placental trisomy 16 mosaicism. *Prenat. Diagn.*, **26**, 956–961.
12. Robinson, W.P., Peñaherrera, M.S., Jiang, R., Avila, L., Sloan, J., McFadden, D.E., Langlois, S. and von Dadelszen, P. (2010) Assessing the role of placental trisomy in preeclampsia and intrauterine growth restriction. *Prenat. Diagn.*, **30**, 1–8.
13. Kalousek, D., Howard-Peebles, P., Olson, S., Barrett, I., Dorfmann, A., Black, S., Schulman, J. and Wilson, R. (1991) Confirmation of CVS mosaicism in term placentae and high frequency of intrauterine growth retardation association with confined placental mosaicism. *Prenat. Diagn.*, **11**, 743–750.
14. Rasmussen, K.M. (1999) Causes of intrauterine growth restriction. *Prenatal Care: Effectiveness and Implementation*, 153–174.
15. Roberts, D.J. and Post, M.D. (2008) The placenta in pre-eclampsia and intrauterine growth restriction. *J. Clin. Pathol.*, **61**, 1254–1260.
16. Benton, S.J., McCowan, L.M., Heazell, A.E.P., Gynspan, D., Hutcheon, J.A., Senger, C., Burke, O., Chan, Y., Harding, J.E., Yockell-Lelièvre, J. et al. (2016) Placental growth factor as a marker of fetal growth restriction caused by placental dysfunction. *Placenta*, **42**, 1–8.
17. Leavey, K., Bainbridge, S.A. and Cox, B.J. (2015) Large scale aggregate microarray analysis reveals three distinct molecular subclasses of human preeclampsia. *PLoS One*, **10**, e0116508.
18. Avila, L., Yuen, R., Diego-Alvarez, D., Peñaherrera, M., Jiang, R. and Robinson, W. (2010) Evaluating DNA methylation and gene expression variability in the human term placenta. *Placenta*, **31**, 1070–1077.
19. Blair, J.D., Yuen, R.K., Lim, B.K., McFadden, D.E., von Dadelszen, P. and Robinson, W.P. (2013) Widespread DNA hypomethylation at gene enhancer regions in placentas associated with early-onset preeclampsia. *Mol. Hum. Reprod.*, **19**, 687–708.
20. Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L., Fan, J.-B. and Shen, R. (2011) High density DNA methylation array with single CpG site resolution. *Genomics*, **98**, 288–295.
21. Blair, J., Langlois, S., McFadden, D. and Robinson, W. (2014) Overlapping DNA methylation profile between placentas with trisomy 16 and early-onset preeclampsia. *Placenta*, **35**, 216–222.
22. Anton, L., Brown, A.G., Bartolomei, M.S. and Elovitz, M.A. (2014) Differential methylation of genes associated with cell adhesion in preeclamptic placentas. *PLoS One*, **9**, e100148.

23. Chu, T., Bunce, K., Shaw, P., Shridhar, V., Althouse, A., Hubel, C. and Peters, D. (2014) Comprehensive analysis of preeclampsia-associated DNA methylation in the placenta. *PLoS One*, **9**, e107318.
24. Jia, R., Zhang, X., Hu, P., Liu, X., Hua, X., Wang, X. and Ding, H. (2012) Screening for differential methylation status in human placenta in preeclampsia using a CpG island plus promoter microarray. *Int. J. Mol. Med.*, **30**, 133.
25. Roifman, M., Choufani, S., Turinsky, A.L., Drewlo, S., Keating, S., Brudno, M., Kingdom, J. and Weksberg, R. (2016) Genome-wide placental DNA methylation analysis of severely growth-discordant monochorionic twins reveals novel epigenetic targets for intrauterine growth restriction. *Clin. Epigenet.*, **8**, 70.
26. Sundrani, D.P., Reddy, U.S., Joshi, A.A., Mehendale, S.S., Chavan-Gautam, P.M., Hardikar, A.A., Chandak, G.R. and Joshi, S.R. (2013) Differential placental methylation and expression of VEGF, FLT-1 and KDR genes in human term and preterm preeclampsia. *Clin. Epigenet.*, **5**, 6.
27. Tobi, E.W., Heijmans, B.T., Kremer, D., Putter, H., Delemarre-van de Waal, H.A., Finken, M.J., Wit, J.M. and Slagboom, P.E. (2011) DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. *Epigenetics*, **6**, 171–176.
28. Banister, C.E., Koestler, D.C., Maccani, M.A., Padbury, J.F., Houseman, E.A. and Marsit, C.J. (2011) Infant growth restriction is associated with distinct patterns of DNA methylation in human placentas. *Epigenetics*, **6**, 920–927.
29. Yuen, R.K., Chen, B., Blair, J.D., Robinson, W.P. and Nelson, D.M. (2013) Hypoxia alters the epigenetic profile in cultured human placental trophoblasts. *Epigenetics*, **8**, 192–202.
30. Yeung, K.R., Chiu, C.L., Pidsley, R., Makris, A., Hennessy, A. and Lind, J.M. (2016) DNA methylation profiles in preeclampsia and healthy control placentas. *Am. J. Physiol. Heart Circ. Physiol.*, **310**, H1295–H1303.
31. Gillis, J., Mistry, M. and Pavlidis, P. (2010) Gene function analysis in complex data sets using ErmineJ. *Nat. Protoc.*, **5**, 1148–1159.
32. Pankov, R. and Yamada, K.M. (2002) Fibronectin at a glance. *J. Cell. Sci.*, **115**, 3861–3863.
33. Wilson, S.L., Blair, J.D., Hogg, K., Langlois, S., von Dadelszen, P. and Robinson, W.P. (2015) Placental DNA methylation at term reflects maternal serum levels of INHA and FN1, but not PAPPa, early in pregnancy. *BMC Med. Genet.*, **16**, 111.
34. Bahr, B.L., Price, M., Merrill, D., Mejia, C., Call, L., Bearss, D. and Arroyo, J. (2014) Different expression of placental pyruvate kinase in normal, preeclamptic and intrauterine growth restriction pregnancies. *Placenta*, **35**, 883–890.
35. Szeverenyi, I., Cassidy, A.J., Chung, C.W., Lee, B.T.K., Common, J.E.A., Ogg, S.C., Chen, H., Sim, S.Y., Goh, W.L.P., Ng, K.W. et al. (2008) The human intermediate filament database: Comprehensive information on a gene family involved in many human diseases. *Hum. Mutat.*, **29**, 351–360.
36. Tan, K.H., Tan, S.S., Sze, S.K., Lee, W.K.R., Ng, M.J. and Lim, S.K. (2014) Plasma biomarker discovery in preeclampsia using a novel differential isolation technology for circulating extracellular vesicles. *Obstet. Gynecol.*, **211**, 380. e1–380. e13.
37. Schluns, K.S., Kieper, W.C., Jameson, S.C. and Lefrançois, L. (2000) Interleukin-7 mediates the homeostasis of naive and memory CD8 T cells in vivo. *Nat. Immunol.*, **1**, 426.
38. Rodriguez, A., Vigorito, E., Clare, S., Warren, M.V., Couttet, P., Soond, D.R., van Dongen, S., Grocock, R.J., Das, P.P., Miska, E.A. et al. (2007) Requirement of bic/microRNA-155 for normal immune function. *Science*, **316**, 608–611.
39. Thapa, M., Welner, R.S., Pelayo, R. and Carr, D.J. (2008) CXCL9 and CXCL10 expression are critical for control of genital herpes simplex virus type 2 infection through mobilization of HSV-specific CTL and NK cells to the nervous system. *J. Immunol.*, **180**, 1098–1106.
40. Powers, R.W., Roberts, J.M., Plymire, D.A., Pucci, D., Datwyler, S.A., Laird, D.M., Sogin, D.C., Jeyabalan, A., Hubel, C.A. and Gandley, R.E. (2012) Low placental growth factor across pregnancy identifies a subset of women with preterm preeclampsia: Type 1 versus type 2 preeclampsia?. *Hypertension*, **60**, 239–246.
41. Wallner, W., Sengenberger, R., Strick, R., Strissel, P.L., Meurer, B., Beckmann, M.W. and Schlembach, D. (2007) Angiogenic growth factors in maternal and fetal serum in pregnancies complicated by intrauterine growth restriction. *Clin. Sci. (Lond)*, **112**, 51–57.
42. Benton, S.J., Hu, Y., Xie, F., Kupfer, K., Lee, S., Magee, L.A. and von Dadelszen, P. (2012) Can placental growth factor in maternal circulation identify fetuses with placental intrauterine growth restriction? *Am. J. Obstet. Gynecol.*, **206**, 163. e1–163. e7.
43. Hauth, J.C., Clifton, R.G., Roberts, J.M., Myatt, L., Spong, C.Y., Leveno, K.J., Varner, M.W., Wapner, R.J., Thorp, J.M., Mercer, B.M. et al. (2012) First-trimester prediction of preeclampsia in nulliparous women at low risk. *Obstet. Gynecol.*, **32**, 1234–1242.
44. Raqib, R., Alam, D.S., Sarker, P., Ahmad, S.M., Ara, G., Yunus, M., Moore, S.E. and Fuchs, G. (2007) Low birth weight is associated with altered immune function in rural bangladeshi children: A birth cohort study. *Am. J. Clin. Nutr.*, **85**, 845–852.
45. Statistics Canada. Ottawa. Released April 2, 2008. (2008) Ethnocultural portrait of Canada highlight tables. 2006 census. Catalogue no. 97-562-XWE2006002.
46. Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E. et al. (2013) Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res.*, **23**, 555–567.
47. Roland, C.S., Hu, J., Ren, C., Chen, H., Li, J., Varvoutis, M.S., Leaphart, L.W., Byck, D.B., Zhu, X. and Jiang, S. (2016) Morphological changes of placental syncytium and their implications for the pathogenesis of preeclampsia. *Cell. Mol. Life Sci.*, **73**, 365–376.
48. Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K. and Kelsey, K.T. (2012) DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, **13**, 86.
49. Houseman, E.A., Molitor, J. and Marsit, C.J. (2014) Reference-free cell mixture adjustments in analysis of DNA methylation data. *Bioinformatics*, **30**, 1431–1439.
50. Edgar, R.D., Jones, M.J., Robinson, W.P. and Kobor, M.S. (2017) An empirically driven data reduction method on the human 450K methylation array to remove tissue specific non-variable CpGs. *Clin. Epigenet.*, **9**, 11.
51. Langfelder, P. and Horvath, S. (2008) WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 559.
52. Peters, T. and Peters, M.T. biocViews DifferentialMethylation G., GenomeAnnotation D., OneChannel T. and MultipleComparison Q. (2016) Package ‘DMRcate’.
53. Irizarry, R.A., Aryee, M., Bravo, H.C., Hansen, K.D., Jaffee, H.A., Irizarry, M.R.A. and RUnit, S. biocViews

- DNAMethylation E. and Infrastructure M. (2013) Package 'bumphunter'.
54. Gluckman, P.D., Buklijas, T. and Hanson, M.A. (2015) The developmental origins of health and disease (DOHaD) concept: Past, present, and future. *The epigenome and developmental origins of health and disease*. Academic, London, 1–13.
  55. Hanson, M.A. and Gluckman, P.D. (2014) Early developmental conditioning of later health and disease: Physiology or pathophysiology?. *Physiol. Rev.*, **94**, 1027–1076.
  56. Magee, L.A., Helewa, M., Rey, E., Cardew, S., Côté, A.-M., Douglas, M.J., Firoz, T., Gibson, P.S., Gruslin A, A., Lange, I. et al. (2008) Hypertension Guideline Committee and Strategic Training Initiative in Research in the Reproductive Health Sciences (STIRRHs) Scholars (2008) Diagnosis, evaluation, and management of the hypertensive disorders of pregnancy. *J. Obstet. Gynaecol. Can.*, **30**, S1–48.
  57. Kramer, M.S., Platt, R.W., Wen, S.W., Joseph, K.S., Allen, A., Abrahamowicz, M., Blondel, B. and Breart, G. and Fetal/Infant Health Study Group of the Canadian Perinatal Surveillance System (2001) A new and improved population-based canadian reference for birth weight for gestational age. *Pediatrics*, **108**, E35.
  58. Buhule, O.D., Minster, R.L., Hawley, N.L., Medvedovic, M., Sun, G., Viali, S., Deka, R., McGarvey, S.T. and Weeks, D.E. (2014) Stratified randomization controls better for batch effects in 450K methylation analysis: A cautionary tale. *Front. Genet.*, **5**, 354.
  59. Leavey, K., Benton, S.J., Grynspan, D., Kingdom, J.C., Bainbridge, S.A. and Cox, B.J. (2016) Unsupervised placental gene expression profiling identifies clinically relevant subclasses of human preeclampsia. *Hypertension*, **68**, 137–147.
  60. American Academy of Pediatrics (2004) National high blood pressure education program working group on high blood pressure in children and adolescents. *Pediatrics*, **114**, 555. iv-iv.
  61. Fortin, J., Labbe, A., Lemire, M., Zanke, B.W., Hudson, T.J., Fertig, E.J., Greenwood, C.M. and Hansen, K.D. (2014) Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol.*, **15**, 1–17.
  62. Price, E.M., Cotton, A.M., Lam, L.L., Farré, P., Emberly, E., Brown, C.J., Robinson, W.P. and Kobor, M.S. (2013) Additional annotation enhances potential for biologically-relevant analysis of the illumina infinium HumanMethylation450 BeadChip array. *Epigenetics Chromatin*, **6**, 1.
  63. Suzuki, R. and Shimodaira, H. (2006) Pvcust: An R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, **22**, 1540–1542.
  64. Huang, H., Liu, Y., Marron, J. and Huang, M.H. (2015) Package 'sigclust'.