Title: Characterization of methylation signatures in spontaneous preterm birth

Heather M Brockway[1*], Samantha L Wilson[2], Suhas G Kallapur[3], Catalin S Buhimschi[4], Louis J Muglia[5], and Helen N Jones[1]

[1]Department of Physiology and Functional Genomics, College of Medicine at the University of Florida, Gainesville, FL, USA

[2]Princess Margaret Cancer Centre, University Health Network, University of Toronto, Toronto ON, Canada

[3]Divisions of Neonatology and Developmental Biology, David Geffen School of Medicine at the University of California, UCLA Mattel Children's Hospital, Los Angeles, CA, USA

[4]Department of Obstetrics and Gynecology, The University of Illinois College of Medicine, Chicago, IL, USA

[5]Burroughs Wellcome Fund, Research Triangle Park, NC, USA

* Corresponding author

Abstract: Preterm birth (PTB) is a global public health crisis which results in significant neonatal and maternal mortality. Yet little is known regarding the molecular mechanisms of idiopathic spontaneous PTB (isPTB) and we have few diagnostic markers for adequate assessment of placental development and function. Previous studies of placental pathology, and our transcriptomics studies suggest a role for placental maturity in isPTB. It is known that placental methylation changes over gestation and we hypothesized that if placental hypermaturity is present in our samples, we would observe unique isPTB methylation signature as well as identify loci where isPTB methylation is more similar to that of term birth (TB) than the gestational age matched controls. Our results indicate the isPTB DNA methylation pattern mimics the TB methylation pattern suggesting hypermaturity. Only seven significant differentially methylated regions (DMRs) fitting the isPTB specific hypomethylation (relative to the controls) pattern were identified, indicating unusually high similarity in DNA methylation between isPTB and TB samples. In contrast, 1,718 acute histologic chorioamnionitis(AHC) specific DMRs were identified with hypermethylated DMRs in WNT and cadherin pathways when compared to isPTB and TB samples. In these AHC

DMRs, there were no significant differences between the isPTB and TB, which indicated again, a striking level of similarly between isPTB and TB sample sets. Taken together, these data reflect a more mature placenta than expected which may be impacting birth timing.

Introduction

Preterm birth, defined as delivery at less than 37 weeks of gestation is the leading cause of neonatal mortality worldwide. Prematurity affects an average of 10% of infants born in the United States with rates increasing and costs approximately $26.2 billion dollars a year (annual societal cost including medical, educational and lost productivity)[1,2]. The majority (50%) of preterm births are idiopathic and spontaneous, rather than being related directly to diagnosed medical causes (e.g. pre-eclampsia). Risk factors include but are not limited to: ethnicity, fetal gender, environmental exposures, and economic disparities[3]. Complications include developmental delays, growth restriction, chronic respiratory problems as well as adult sequale[3]. Studies into the etiology of preterm birth have implicated a role for the placenta, a central component of the maternal-fetal interface, which has a vital role in pregnancy initiation, maintenance, and birth timing as well as fetal growth and development[4]. As such, proper placental development, maturation, and function are essential for a successful pregnancy outcome and life time offspring health. Each of these processes is an intricate balance of molecular interactions that are not fully understood even in healthy, normal, term pregnancies. Placental maturation is accompanied by a marked increase in placental surface area ($6m^2$ to $12m^2$) due to placental remodeling initiated between 20-24 weeks gestation and continuing throughout the remainder of gestation[4]. This remodeling accommodates exponential fetal growth across the second half of gestation. Under normal physiological conditions, placental maturation is recognized by histological hallmarks including increased quantities of terminal villi (<80 microns in diameter), syncytial nuclear aggregates (SNAs, 10+ syncytial nuclei being extruded from the syncytiotrophoblast), and formation of the vasculosyncytial membranes (VSM) which allow for more efficient transport of nutrients and diffusion of blood gases across the syncytiotrophoblast and villous vascular endothelial

cell membranes to the placental-fetal circulation[5,6]. When these hallmarks are observed in significant quantities prior to 37 weeks, placentas are classified as having advanced villous maturation (AVM). Histological studies of pathological placentas indicate AVM occurs in 50-60% of isPTB and iatrogenic preterm births[7,8]. This indicates a potential developmental disconnect between placental maturation and the corresponding fetal maturation. In infection associated preterm births, AVM was observed in less than 20% of pathologic placentas[7,8]. These studies indicate multiple morphological endotypes exist, underlying the classical clinical PTB phenotypes, especially those of spontaneous PTB which are based on gestational age and simply defined as early, moderate and late[9]. The identification of these morphological endotypes further highlights the heterogeneity confounding the identification of PTB etiology and potential diagnostic biomarkers. Multiple levels of heterogeneity confound elucidation of molecular mechanisms involved in sPTB, from inconsistent sampling of maternal/placental/fetal tissues to the numerous cell types across the maternal-fetal interface[10–13]. One way to overcome this problem is to link morphological endotypes to well characterized molecular signatures from the same sample to generate more precise spontaneous preterm birth phenotypes.

Over the last decade, advances in the integration of "omics" data have allowed for the discovery of biomarkers and mechanistic insight into various diseases including several types of cancer[14–16]. However, our limited knowledge of normal placental physiology throughout gestation coupled with a lack of precise preterm morphological and molecular phenotyping impedes our ability to exploit omics" data to improve pregnancy outcomes associated with prematurity. While transcriptomics allows for examination of differential gene expression between normal and PTB placental tissues, DNA methylation (DNAm) studies on the same tissue may provide insights into the observed differences through regulation of expression[17–20]. However, heterogeneity of DNAm across placental villous tissues has made interpretation of such data challenging[19]. To further confound interpretation, some studies have focused on DNAm differences between normal and iatrogenic preterm

conditions such as fetal growth restriction (FGR), pre-eclampsia (PE), or simply differences in gestational age, or PTB as a whole without consideration for underlying pathophysiology such as placental maturation[21–23].

We have previously identified transcriptomic signatures of AVM in a small cohort using clinically phenotyped placental villous samples from isPTB and AHC births between 29 and 36 weeks and normal term births between 38 and 42 weeks[24]. Given the importance of DNA methylation to placental development and maturation, we wanted determine if we could detect DNAm signatures associated with spontaneous PTB phenotypes. As with our previous transcriptomic analyses, we were able to identify distinct DNAm signatures indicative of AVM in our isPTB and AHC samples.

Methods

Study Population

This study was approved by the Cincinnati Children's Hospital Medical Center institutional review board (#IRB 2013-2243, 2015-8030, 2016-2033). De-identified term (n=6), isPTB (n=8), and IAI (n=8) placental villous samples along with appropriate covariate information were obtained from the following sources: The Global Alliance to Prevent Prematurity and Stillbirth (GAPPS) in Seattle Washington USA, the Research Centre for Women's and Infant's Health (RCWIH) at Mt Sinai Hospital Toronto Canada, and the University of Cincinnati Medical Center (UCMC). Inclusion criteria included: maternal age 18 years or older, singleton pregnancies with either normal term delivery (38-42 weeks gestation) or preterm delivery (29-36 weeks gestation) without additional complications. Additional data regarding these samples can be found in[24].

DNA Methylome Generation

DNA was isolated from homogenized, snap frozen placental villous samples using the DNAeasy Kit (Qiagen). DNA quality was assessed using Qubit and nanodrop. A minimum of 500ng was submitted to the University of Minnesota Genomics Center and the University of Cincinnati Genomics, Epigenomics and Sequencing Core for

DNA quality assessment, bisulfite conversion, and methylome generation on the Illumina Methylation EPIC Bead Chip. Control samples were included in each run to assess for batch effects across different array chips.

DNA Methylation array data processing

Methylation data processing and analyses based on a previously developed workflow[25] which was modified to fit the analysis parameters of this study. All packages are available within Bioconductor[26] and all package scripts were run in RStudio/R v4.0.2[27,28]. IDAT file preprocessing and probe quality control was conducted in R using scripts from which are based in *minfi*[29] and *methylumi*[30]. IDAT files and a sample file containing covariate and BeadChip metadata were loaded into R where data quality was assessed using the mean detection p-values for the probes in each sample. This summary data allowed determination of failed samples to be excluded. We assessed which normalization algorithm would best suit our data set, comparing several methods including quantile[31], subset-quartile within array normalization (SWAN)[32], beta-mixture quantile normalization BMIQ[33] and functional normalization(Funnorm)[34]. We chose to apply Functional Normalization (preprocessFunnorm) for the algorithm's ability to utilize the internal control probes for each individual sample in an unsupervised manner to control for unwanted variation.

After normalization, we excluded individual low-quality probes with a detection p-value > 0.1 in more than 2 samples or bead count <3 in at least 5% of samples, sex chromosome probes, cross-hybridizing probes, and probes were SNPs (within the binding region or within 5-10bp of the binding region) could potentially affect hybridization[25]. To ensure appropriate filtering of problematic probes, we utilized several resources including the Illumina MethylationEPIC BeadChip hg38 manifest and Zhou et al[35] to identify additional variation that would interfere with probe hybridization. We utilized McCartney et al[36] to filter the cross-hybridizing probes that are not listed in the manifest. We removed all probes that reside in the ENCODE DAC black list regions[37].

Once probe filtering was complete, we assessed the data for batch effects using principle component analysis (PCA) and no significant batch effect was observed, therefore no correction was applied[38]. The resulting

data matrix contained M-values and was used in all downstream analyses for maximum statistical inference in calculation of differential methylation.

Identification of differentially methylated positions

To assess differentially methylated positions (DMPs), we utilized a generalized linear model (glm) within *limma*[39] to assess differential methylation for each individual probe within the M-value matrix as [25] with adjustment for birth types and fetal sex within the model. We did not assess any additional covariate data in this particular analysis. The following pairwise comparisons were used to identify significant positions of differential methylation: isPTB verses AHC, TB verses AHC and isPTB verses TB. Within *limma*, we utilized the separate method, which applies multiple test adjustments to each column of p-values of each individual pairwise comparison. Multiple corrections testing was performed using the Benjamini Hochberg method[40] using multiple Q values: <0.5, <0.1, <0.2 and <0.3 (Supplemental Table 1). We opted to define significant DMPs with a Q <0.3 and a log2 fold-change of >±1.

Methylome DMP Signature Identification

To identify methylation signatures, we used Venny 2.0[41] to generate Venn diagrams to identify and sort significant DMPs. Candidate DMP fold changes were sorted by methylation pattern. The isPTB signature was defined as any DMP that was more or less methylated when compared to the AHC and TB, with the AHC vs TB methylation signature being non-significant. The AHC signature was defines as any DMP that was more or less methylated compared to isPTB and TB and where the isPTB vs TB methylation was non-significant. Heatmaps were generated in Prism v8 (GraphPad) using M-values. To assess if the differential methylation was influenced by outliers or by artifacts, we generated violin plots with median and quartiles in Prism v8 to check the distribution of the individual sample beta values. Beta values were generated by transforming the M-values from the data matrix.

Differentially Methylation Region (DMR) Identification

We used *DMRcate*[25,42] to identify differentially methylated regions comprised of significant DMPs within a specified distance using moderated t statistics. To identify significant DMPs within *DMRcate*, we used the M-value matrix (normalized and filtered) and set a threshold of Benjamini Hochberg adjusted p-value <0.3. Since *DMRcate* uses *limma* to determine the significant DMPs, we were able to utilize the same glm design from the initial DMPs analysis against adjusting for fetal sex and birth type. Once significant DMPs were identified, DMR identification thresholds were set at lamba=1000, C=2, and minimum cpgs=5. As we are analyzing array data, we opted to use the default lambda and C (scaling factor) which allows for optimal differentiation with 1 standard deviation of support to account for Type 1 errors. Once significant DMRs were identified in each pairwise comparison, we intersected them using Venny 2.0 to identify isPTB and AHC specific DMRs. The isPTB signature was defined as any DMR that was more or less methylated when compared to the AHC and TB, with the AHC vs TB. The AHC signature was defines as any DMR that was more or less methylated compared to isPTB and TB and where the isPTB vs TB methylation was non-significant meaning no DMR was identified in *DMRcate*. We also set a mean difference in differentiation threshold of 0.01. Heatmaps were generated in Prism v8 (GraphPad) using M-values.

Functional analyses of DMRs

DMRs were entered into the Panther Pathway DB[43] for statistical overrepresentation analyses for Reactome Pathways and to assess gene ontology (GO) for biological and molecular processes. Fisher's Exact tests were used to determine significance and Bonferroni correction for multiple comparisons. Pathways were considered significant if they had an adjusted p-value <0.05.

Statistical Analyses

Cohort data were analyzed in Prism v8 (GraphPad). Data were evaluated for normality and non-parametric tests applied as appropriate. Non-parametric data are expressed as median and range and were analyzed by Kruskal-Wallis Test ANOVA with Dunn's Multiple Comparisons. Categorical data were analyzed using Fisher's Exact Test.

Results

Methylation Study Characteristics

Maternal and fetal characteristics for the three different pregnancy outcomes included in the DNAm analyses are presented in Table 1. Transcriptomes from these samples were previously published[24]. Due to the amount of sample required for DNA extraction only a subset of the samples were used. Significant differences were observed in gestational age and fetal weights between AHC and isPTB samples compared to the TB samples ($P<0.05$). All AHC and TB for which there were fetal weights available were appropriate for gestational age. We included males and females in each sample set and adjusted the linear models for fetal sex in addition to birth outcome. It is important to note that in this study, we have mixed ethnic background within each of the sample sets.

Identification of significant differentially methylated positions (DMP)

Preliminary quality control identified one sample with mean probe detection p value >0.1 and it was subsequently removed from methylation analyses. Prior to normalization and subsequent probe filtering, there were 866,901 probes in the data matrix. After normalization and filtering, 108,691 probes were removed, leaving 758,210 probes in the matrix for analyses (Supplemental Table 1).

Our initial statistical testing using the Benjamini Hochberg Q cutoff of 0.05 did not yield any significant DMPs in the isPTB vs TB pairwise comparison. We ran several different cutoffs and ultimately relaxed our Q cutoff <0.3 to obtain a viable number of DMPs in the isPTB vs TB pairwise comparison (Supplemental Table 2). We then set a threshold for differential methylation of log2 fold change of >1. The DMP analysis identified a

total of 24,202 significant DMPs across all pairwise comparisons in the model. In the isPTB vs AHC comparison we identified 8,309 DMPs, 4,334 with reduced methylation and 3,975 more methylated in isPTB compared to AHC. In the TB vs AHC comparison, we identified a total of 15,817 DMPs with 7,170 less methylated and 8,647 more methylated in TB. Lastly, in the isPTB vs TB comparison, 85 DMPs were identified as significant with 13 more methylated and 72 less methylated (Figure 1A).

We observed differences in genomic location of the DMPs between the pairwise comparisons and thus, analyzed the genomic location distribution of the DMPs per comparison (Figure 1B). In the isPTB vs AHC and TB vs AHC comparisons the majority of DMPs were associated with CpG islands, shores, shelves (isPTB = 70% and TB = 65%) while the remaining DMPs were in open sea locations which are typically 3-4kb away from CpG islands (isPTB = 30% and TB = 35% respectively). In contrast, in the isPTB vs TB comparison, 70% of the DMPs were associated with open sea positions while only 30% associated with CpG islands, shores, and shelves.

Isolation of isPTB and AHC DNA methylation signatures using DMPs

The first step in identification of a DMP methylation signature was to intersect the significant DMPs from each pairwise comparison and determine which would potentially segregate into an isPTB or AHC signature (Figure 1C). As a result of the intersection, we identified 47 potential isPTB specific DMPs. Upon examining the DNAm patterns for these DMPs across all pairwise comparisons, we ultimately isolated 3 isPTB specific DMPs. Our examination of the individual sample beta values and their distribution for each DMP to confirm our findings are not due to artifacts or outliers(Figure 2A). Although we initially identified 8,306 potential AHC specific DMPs via the intersection, upon further examination of the DNAm pattern, we ultimately isolated 6,177 (Figure 2B). Of these, 3,002 are more methylated and 3,175 are less methylated. We also examined the genomic location distribution of the AHC signature DMPs and found that 76% were located within CpG islands, shores, and shelves with remaining 24% located in open sea regions (Supplemental Figure 1).

Identification of differentially methylated regions (DMRs)

To identify differentially methylated regions, we used the M-value matrix of data values previously generated in our initial analyses. We utilized again a relaxed Q <0.3 to ensure we would be able to identify enough CpG sites to identify DMRs in the isPTB vs TB comparison (Supplemental Table 5). Only then, we were able to identify significant DMRs within all pairwise comparisons (Table 2). 56 DMRs were observed within the isPTB vs TB comparison in contrast to the thousands significant DMRs identified in the isPTB and TB verses AHC pairwise comparisons. All isPTB vs TB DMRs were under 2000bp wide and had no more than 18 CpG sites in any given DMR. In contrast, the DMRs in the isPTB and TB vs AHC comparisons were wider and encompassed more probes (Table 2). We intersected the DMRs and identified potential candidate DMRs for isPTB and AHC methylation signatures (Supplemental Figure 2). Ultimately, we identified 51 potential isPTB specific and 12,843 AHC specific DMRs. These DMRs overlap with coding and non-coding loci across the genome as per the annotation from *DMRcate* package[42].

Identification and function of DMRs specific to isPTB and AHC

Of the 51 candidate isPTB DMRs, only seven demonstrated an isPTB specific signature (Figure 4 and Table 3). Six isPTB specific DMRs overlap coding/non-coding loci with only one sitting in an upstream promoter region, *LINC02028* (Table 4*)*. This is the only isPTB-specific DMR that overlaps with a CpG island. Four of the DMRs sit within transcripts for *FAM186A*, *NOD2, UBL7-AS1,* and *PDE9A*, more specifically within introns or at intron/exon boundaries. The remaining two DMRs sit in the 3'UTR of genes, *ZBTB4* and *STXB6,* with the *ZBTB4* DMR crossing the last exon/UTR boundary (Table 4). No over-represented pathways were identified.

Of the 12,843 AHC specific DMRs, only 1,718 demonstrated an AHC specific methylation pattern. These DMRs include coding and non-coding loci (Figure 5A and Supplemental Table 6). Of these, 801 DMRs are more methylated while 917 are less methylated than corresponding DMRs in the isPTB or TB pairwise comparison. In the top 25 more/less methylated loci, the lack of significant differences in methylation can be clearly be observed in TB vs isPTB(Figure 5B and Table 5).

We assessed the potential implications of the AHC specific DMRs using statistical over-representation analyses for pathways and GO terms. In the more methylated DMRs, we identified two significantly over-represented pathways: WNT and Cadherin signaling (Table 6). Significant Biological Process GO terms included homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156) and cell-cell adhesion via plasma-membrane adhesion molecules (GO:0098742).

No significant over-represented pathways were identified in the less methylated DMRs. The significant Biological Process GO terms that were associated with the less methylated dataset include cell morphogenesis involved in differentiation (GO:0000904), cell morphogenesis (GO:0000902), cell morphogenesis (GO:0000902), and detection of chemical stimulus (GO:0009593). For Molecular Function, the following significant GO terms were identified: ion binding (GO:0043167), protein binding (GO:0005515), protein binding (GO:0005515), and olfactory receptor activity (GO:0004984) (Table 7)

Discussion

To gain insight into the role of DNA methylation in spontaneous preterm birth, we utilized pairwise comparisons of placental villous tissue from spontaneous preterm births and normal term births within a general linear model adjusting for fetal sex and gestational age at delivery. We were able to identify distinct methylation signatures at both the positional (DMP) and regional (DMR) levels in isPTB and AHC. Through bioinformatic functional assessment, we were able to identify pathways of interest pertaining to placental maturation.

Given the sheer number of datapoints being examined, we felt that relaxing the Q value to 0.3 would not adversely affect our analyses and we were willing to accept the potential increase in false positives[44,45]. This allowed us to better assess any potential differences between isPTB and TB despite the potential increase in false positives. The Benjamini Hochberg correction is dependent on the overall number of samples to be corrected and considered to be rather conservative. Regardless of the statistical parameters applied, the isPTB signature mimicked the TB signature to a high degree which is in agreement with the transcriptomic signatures we

previously identified[24] and provides additional evidence of a potential placental hypermaturity signature associated with isPTB. Although this is the first study investigating DNA methylation in spontaneous preterm birth, this pattern of DNA methylation was also observed in studies of iatrogenic preterm births in DMP and DMR analyses, for both PE and IUGR[20]. In the second study, focusing on imprinted regions found that IUGR samples also mimicked the PE and term controls[46]. Pyrosequencing from this second study confirmed no differences in the DMRs suggesting the detection of hypermaturity molecular signature. Given that hypermaturity is estimated to affect 50-60% of all preterm births[7,8], these results provide additional evidence supporting that DNAm could be clinically useful in classifying various PTB placental pathophysiologies such as hypermaturity[20,47].

DMRs are associated with numerous disease pathologies including various types of tissues[48,49]. While DNAm has been studied in the other adverse pregnancy outcomes such as PE, IUGR, this study is the first to look specifically at isPTB. Our analysis resulted in the identification of seven DMRs with isPTB specific methylation patterns; two are associated with non-coding transcripts (*LINC02028* and *UBL7-AS),* five with genes (*ZBTB4, STXBP6, PDE9A, NOD2,* and *FAM186A).* Of these genes, four are of particular interest due to their potential function in or previous association with PTB.

ZBTB4 is a placentally expressed gene coding for a transcription factor that binds methylated CpGs in a repressive manner, controls TP53 responses in cells, and inhibits cell growth and proliferation [50–52]. TP53 [35]was identified as a potential biological pathway of interest in our microarray meta-analysis of spontaneous PTB[53] and has been implicated in isPTB from a uterine perspective in mice[54]. *STXBP6,* also known as *AMISYN,* binds SNARE complex proteins together[55]. As SNARE complexes have been well described in synaptic vesicle formation and exocytosis[56] and regulation of membrane fusion dynamics[57,58], the presence of this protein in the placenta suggests potential role in placental extracellular vesicle formation or the mediation of membrane fusion during cytotrophoblast differentiation[57,59].

PDE9A functions in the hydrolysis of cAMP into monophosphates, modulating the bioavailability of cAMP and cGMP in cells[60]. cAMP signaling is essential to cytotrophoblast differentiation into syncytiotrophoblast[61]; therefore, alteration of *PDE9A* expression or function impacts cAMP bioavailability potentially altering this specific trophoblast differentiation pathway. In fact, *PDE9A* has been proposed as a potential first trimester maternal serum biomarker for Trisomy 21[62]. Placentas from Trisomy 21 fetuses have multiple defects in cytotrophoblast differentiation, specifically cell fusion, resulting in what appears to be delayed villous maturation, indicating a key role for this gene in normal placental maturation[62–65].

NOD2 has a role in activation of the innate inflammatory response and has been implicated in NFKB activation[66–68]. NFKB activation is a central component of pro-inflammatory /labor pathways in both normal term and preterm pathophysiology[67,69,70]. As a member of the NOD-like receptor family, NOD2 has been previously associated with recognition of pathogen associated molecular patterns (PAMPs) and damage associated molecular patterns (DAMPs) both of which have been associated with preterm labor and birth[67]. The activation of pathways associated with PAMPs and DAMPs have previously been associated with sPTB and iatrogenic PTB[53,71–73]. *NOD2* has been studied primarily in the context of a proinflammatory factor in fetal membranes and myometrium; however, *NOD2* is expressed in first trimester and term placental tissues, specifically in syncytiotrophoblast and stromal cells[66,74]. Furthermore, *NOD2* polymorphisms have been associated with preterm birth in several genetic studies examining innate immunity, preterm premature rupture of membranes (PPROM), and early onset PE and HELLP syndromes[67,72,75,76].

Taken together, these isPTB DMRs and their associated genes suggest that altered DNA methylation maybe highly influential in isPTB; however, from these data alone, it cannot be determined if this is a causative effect or the result of isPTB as the samples were obtained at delivery. Although we cannot sample placental tissues throughout gestation to determine cause or effect, using DNAm profiling on delivered placental tissues will provide key insights in the pathophysiological underpinnings of adverse pregnancy outcomes.

In contrast to the isPTB DNAm signature, our examination of the AHC samples compared to the isPTB and TB samples identified 1,718 DMRs. We observed within the top 25 more/less methylated DMRs, multiple DMRs were associated with genes of interest that were previously associated with adverse pregnancy outcomes including IUGR and PE. Several have also been associated gestational diabetes mellitus (GDM) which can also result in preterm birth. These genes of interest include: *MLLT1*[77], *FGFR2*[77], *CACNA1A*[78], *GCK*[79,80], *FER1L6*[81], *CTSH*[82], and *ACAP3*[83]. Additionally, *GSE1* [84], *VSTM1*[85], and *ACSS1*[84] are expressed in the placenta but have not yet been associated with an adverse pregnancy outcome. Our pathway analyses of the more methylated DMRs, yielded two pathways with statistical over-representation, WNT and Cadherin signaling. Both of these pathways are necessary for placental development and maturation[86–89] and a prior methylation study in PE also identified differential methylation (increased methylation) in WNT and cadherin signaling[90], which is in agreement with our findings. Given that over 50% of PE cases have hypermaturity along with the pathological hallmarks of PE, this may indicate a role for these pathways in placental maturation.

One of the caveats to studying placental villous omics of any nature is the lack of normal gestational age matched tissue due to limited accessibility throughout gestation. We previously utilized infection associated samples in our transcriptome analyses as our gestational age controls as their villi did not appear to be inflamed via pathological assessment. While we cannot rule out that changes at AHC loci may be due to infection, we did not observe pathways or GO terms associated with immunity or infection. Our data suggests that the overall AHC DNAm signature is reflective of appropriate villous maturation rather than an infection signature as was observed in our transcriptome data[24].

This is the first study to examine DNAm in spontaneous preterm birth in the context of placental maturity. The identification of hypermaturity signatures by both positional and regional differences in methylation highlights the necessity of spontaneous preterm placentas in order to understand underlying the molecular mechanisms leading to the observed hypermaturity. These differences could be due to altered trophoblast biology.

These data when taken in the context of a potential epigenetic clock, suggests that perhaps epigenetic aging may have a role as it has in other fetal tissue and stem cells[91,92]. Future studies need to investigate the origin of the observed hypermaturity and its impact on the maternal-fetal interface and pregnancy outcomes.

Acknowledgements: The authors would like to express their gratitude to the patients who donated their placentas for research. We would also like to thank Pietro Presicce, Paranthaman Senthamarai Kannan, and Manuel Alvarez (Kallapur lab), William E. Ackerman, Irina A. Buhimschi (Buhimschi lab), GAPPS, and RCWIH for assisting us in obtaining the placental samples and covariate data. Additionally, the authors thank the staff at the University of Cincinnati Genomics, Epigenomics and Sequencing Core and the University of Minnesota Genomics Center for their assistance in generating the methylation data for this project. This work was supported by an Interdisciplinary Scholar Award to HMB as part of the March of Dimes Prematurity Research Center Ohio Collaborative Funding 07/136/18 and LJM, HNJ, and HMB were supported by the Eunice Kennedy Shriver National Institute of Child and Health & Human Development of the National Institutes of Health Award Number R01HD091527. While the funders provided salary support for three authors, they had no further role in the study design, data collection and analyses, decision to publish or preparation of this manuscript.

Author Contributions: Conceptualization: Heather M. Brockway. Data curation: Heather M. Brockway. Formal analysis: Heather M. Brockway. Funding acquisition: Heather M. Brockway, Louis J. Muglia and. Helen Jones. Project administration: Heather M. Brockway. Resources: Samantha L. Wilson, Suhas G. Kallapur, Catalin S. Buhimschi, Louis J. Muglia, Helen N. Jones. Supervision: Louis J. Muglia, Helen N. Jones. Data Interpretation: Heather M. Brockway and Samantha L. Wilson, and Helen N Jones. Writing – original draft: Heather M. Brockway. Writing – review & editing: Heather M. Brockway, Samantha L. Wilson, Suhas G. Kallapur, Catalin S. Buhimschi, Louis J. Muglia, Helen N. Jones.

Disclosure of Interest Statement: The authors report no conflict of interest for this research supported by National Institutes of Health Award Number R01HD091527

Data statement: Data will be uploaded and released in the NIH Gene Expression Omnibus upon final acceptance of this manuscript as per funding agency guidelines.

References:

1. Blencowe H, Cousens S, Chou D, Oestergaard M, Say L, Moller A-B, Kinney M, Lawn J. Born Too Soon: The global epidemiology of 15 million preterm births. Reproductive Health 2013; 10:S2.

2. Chawanpaiboon S, Vogel JP, Moller A-B, Lumbiganon P, Petzold M, Hogan D, Landoulsi S, Jampathong N, Kongwattanakul K, Laopaiboon M, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. The Lancet Global Health 2018; 7:e37–46.

3. Monangi NK, Brockway HM, House M, Zhang G, Muglia LJ. The genetics of preterm birth: Progress and promise. Seminars in perinatology 2015; 39:574–83.

4. Burton GJ, Fowden AL. The placenta: a multifaceted, transient organ. Philosophical Transactions of the Royal Society of London B: Biological Sciences 2015; 370:20140066.

5. Khong TY, Mooney EE, Ariel I, Balmus NCM, Boyd TK, Brundler M-A, Derricott H, Evans MJ, Faye-Petersen OM, Gillan JE, et al. Sampling and Definitions of Placental Lesions: Amsterdam Placental Workshop Group Consensus Statement. Archives of Pathology & Laboratory Medicine 2016; 140:698–713.

6. Sankar DK, Bhanu SP, Kiran S, Ramakrishna B, Shanthi V. Vasculosyncytial membrane in relation to syncytial knots complicates the placenta in preeclampsia: a histomorphometrical study. Anatomy & Cell Biology 2012; 45:86–91.

7. Morgan TK. Role of the Placenta in Preterm Birth: A Review. American Journal of Perinatology 2016; 33:258–66.

8. Nijman TA, Vliet EO van, Benders MJ, Mol BW, Franx A, Nikkels PG, Oudijk MA. Placental histology in spontaneous and indicated preterm birth: A case control study. Placenta 2016; 48:56–62.

9. Manuck TA, Esplin SM, Biggio J, Bukowski R, Parry S, Zhang H, Huang H, Varner MW, Andrews W, Saade G, et al. The phenotype of spontaneous preterm birth: application of a clinical phenotyping tool. American Journal of Obstetrics and Gynecology 2015; 212:487.e1-487.e11.

10. Benton SJ, Leavey K, Grynspan D, Cox BJ, Bainbridge SA. The clinical heterogeneity of preeclampsia is related to both placental gene expression and placental histopathology. American journal of obstetrics and gynecology 2018; 219:604.e1-604.e25.

11. Moutquin J-M. Classification and heterogeneity of preterm birth. BJOG: An International Journal of Obstetrics and Gynaecology 2003; 110:30–3.

12. Yuen RKC, Robinson WP. Review: A high capacity of the human placenta for genetic and epigenetic variation: Implications for assessing pregnancy outcome. Placenta 2011; 32:S136–41.

13. Whigham C-AA, MacDonald TM, Walker SP, Hannan NJ, Tong S, Kaitu'u-Lino TJ. The untapped potential of placenta-enriched molecules for diagnostic and therapeutic development. Placenta 2019;

14. Benny PA, Al-Akwaa FM, Schlueter RJ, Lassiter CB, Garmire LX. A review of omics approaches to study preeclampsia. Placenta 2020; 92:17–27.

15. Chabrun F, Huetz N, Dieu X, Rousseau G, Bouzillé G, Barca JMC de la, Procaccio V, Lenaers G, Blanchet O, Legendre G, et al. Data-Mining Approach on Transcriptomics and Methylomics Placental Analysis Highlights Genes in Fetal Growth Restriction. Frontiers in Genetics 2020; 10:1292.

16. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. Genome Biology 2017; 18:83.

17. Yuen RK, Peñaherrera MS, Dadelszen P von, McFadden DE, Robinson WP. DNA methylation profiling of human placentas reveals promoter hypomethylation of multiple genes in early-onset preeclampsia. European journal of human genetics : EJHG 2010; 18:1006–12.

18. Parets SE, Conneely KN, Kilaru V, Menon R, Smith AK. DNA methylation provides insight into intergenerational risk for preterm birth in African Americans. Epigenetics 2015; 10:784–92.

19. Avila L, Yuen RK, Diego-Alvarez D, Peñaherrera MS, Jiang R, Robinson WP. Evaluating DNA methylation and gene expression variability in the human term placenta. Placenta 2010; 31:1070–7.

20. Wilson SL, Leavey K, Cox BJ, Robinson WP. Mining DNA methylation alterations towards a classification of placental pathologies. Human molecular genetics 2018; 27:135–46.

21. Lim Y, Li J, Ni Y, Liang Q, Zhang J, Yeo GS, Lyu J, Jin S, Ding C. A complex association between DNA methylation and gene expression in human placenta at first and third trimesters. PLOS ONE 2017; 12:e0181155.

22. Mayne BT, Leemaqz SY, Smith AK, Breen J, Roberts CT, Bianco-Miotto T. Accelerated placental aging in early onset preeclampsia pregnancies identified by DNA methylation. Epigenomics-uk 2016; 9:279–89.

23. Toure DM, ElRayes W, Barnes-Josiah D, Hartman T, Klinkebiel D, Baccaglini L. Epigenetic modifications of human placenta associated with preterm birth: a systematic review. The Journal of Maternal-Fetal & Neonatal Medicine 2017; 31:1–12.

24. Brockway HM, Kallapur SG, Buhimschi IA, Buhimschi CS, Ackerman WE, Muglia LJ, Jones HN. Unique transcriptomic landscapes identified in idiopathic spontaneous and infection related preterm births compared to normal term births. PloS one 2019; 14:e0225062.

25. Maksimovic J, Phipson B, Oshlack A. A cross-package Bioconductor workflow for analysing methylation array data. F1000Research 2016; 5:1281.

26. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, et al. Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods 2015; 12:115–21.

27. Team Rs. RStudio: Integrated Development for R. RStudio. 2020; Available from: http://www.rstudio.com/.

28. Team RC. R: A language and environment for statistical ## computing. R Foundation for Statistical Computing [Internet] 2020; Available from: https://www.R-project.org/

29. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 2014; 30:1363–9.

30. S D, P D, S B, Triche, T Jr, M B. methylumi: Handle Illumina methylation data. 2020; Available from: https://www.bioconductor.org/packages/release/bioc/html/methylumi.html

31. Wang T, Guan W, Lin J, Boutaoui N, Canino G, Luo J, Celedón JC, Chen W. A systematic study of normalization methods for Infinium 450K methylation data using whole-genome bisulfite sequencing data. Epigenetics 2015; 10:662–9.

32. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile Within Array Normalization for Illumina Infinium HumanMethylation450 BeadChips. Genome Biol 2012; 13:R44.

33. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics 2013; 29:189–96.

34. Fortin J-P, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, Greenwood CM, Hansen KD. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biology 2014; 15:503.

35. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Research 2017; 45:e22–e22.

36. McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. Genomics Data 2016; 9:22–4.

37. Amemiya HM, Kundaje A, Boyle AP. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Scientific Reports 2019; 9:9354.

38. Price EM, Robinson WP. Adjusting for Batch Effects in DNA Methylation Microarray Data, a Lesson Learned. Frontiers in genetics 2018; 9:83.

39. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research 2015; 43:e47–e47.

40. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 1995; 57:289–300.

41. Oliveros, J.C. Venny. An interactive tool for comparing lists with Venn's diagrams. 2015; Available from: https://bioinfogp.cnb.csic.es/tools/venny/index.html

42. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, Lord RV, Clark SJ, Molloy PL. De novo identification of differentially methylated regions in the human genome. Epigenetics & Chromatin 2015; 8:6.

43. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Research 2017; 45:D183–9.

44. Li D, Xie Z, Pape M, Dye T. An evaluation of statistical methods for DNA methylation microarray data analysis. BMC Bioinformatics 2015; 16:217.

45. Mansell G, Gorrie-Stone TJ, Bao Y, Kumari M, Schalkwyk LS, Mill J, Hannon E. Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. Bmc Genomics 2019; 20:366.

46. Monteagudo-Sánchez A, Sánchez-Delgado M, Mora JRH, Santamaría NT, Gratacós E, Esteller M, Heredia ML de, Nunes V, Choux C, Fauque P, et al. Differences in expression rather than methylation at placenta-specific imprinted loci is associated with intrauterine growth restriction. Clin Epigenetics 2019; 11:35.

47. Wilson SL, Robinson WP. Utility of DNA methylation to assess placental health. Placenta 2018; 64 Suppl 1:S23–8.

48. Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. Biochimica Et Biophysica Acta Bba - Rev Cancer 2007; 1775:138–62.

49. Ehrlich M. DNA hypermethylation in disease: mechanisms and clinical relevance. Epigenetics 2019; 14:1–23.

50. Filion GJP, Zhenilo S, Salozhin S, Yamada D, Prokhortchouk E, Defossez P-A. A Family of Human Zinc Finger Proteins That Bind Methylated DNA and Repress Transcription. Mol Cell Biol 2006; 26:169–81.

51. Yu Y, Shang R, Chen Y, Li J, Liang Z, Hu J, Liu K, Chen C. Tumor suppressive ZBTB4 inhibits cell growth by regulating cell cycle progression and apoptosis in Ewing sarcoma. Biomed Pharmacother 2018; 100:108–15.

52. Weber A, Marquardt J, Elzi D, Forster N, Starke S, Glaum A, Yamada D, Defossez P, Delrow J, Eisenman RN, et al. Zbtb4 represses transcription of P21CIP1 and controls the cellular response to p53 activation. Embo J 2008; 27:1563–74.

53. Paquette AG, Brockway HM, Price ND, Muglia LJ. Comparative transcriptomic analysis of human placentae at term and preterm delivery. Biology of reproduction 2018; 98:89–101.

54. Hirota Y, Daikoku T, Tranguch S, Xie H, Bradshaw HB, Dey SK. Uterine-specific p53 deficiency confers premature uterine senescence and promotes preterm birth in mice. Journal of Clinical Investigation 2010; 120:803–15.

55. Scales SJ, Hesser BA, Masuda ES, Scheller RH. Amisyn, a Novel Syntaxin-binding Protein That May Regulate SNARE Complex Assembly*. J Biol Chem 2002; 277:28271–9.

56. Wang T, Li L, Hong W. SNARE proteins in membrane trafficking. Traffic 2017; 18:767–75.

57. Han J, Pluhackova K, Böckmann RA. The Multifaceted Role of SNARE Proteins in Membrane Fusion. Front Physiol 2017; 8:5.

58. Bogaart G van den, Jahn R. Counting the SNAREs needed for membrane fusion. J Mol Cell Biol 2011; 3:204–5.

59. Guček A, Gandasi NR, Omar-Hmeadi M, Bakke M, Døskeland SO, Tengholm A, Barg S. Fusion pore regulation by cAMP/Epac2 controls cargo release during insulin exocytosis. Elife 2019; 8:e41711.

60. Rentero C, Puigdomènech P. Specific use of start codons and cellular localization of splice variants of human phosphodiesterase 9A gene. Bmc Mol Biol 2006; 7:39.

61. Gerbaud P, Taskén K, Pidoux G. Spatiotemporal regulation of cAMP signaling controls the human trophoblast fusion. Front Pharmacol 2015; 6:202.

62. Lim JH, Kim SY, Park SY, Lee SY, Kim MJ, Han YJ, Lee SW, Chung JH, Kim MY, Yang JH, et al. Non-Invasive Epigenetic Detection of Fetal Trisomy 21 in First Trimester Maternal Plasma. Plos One 2011; 6:e27709.

63. Frendo JL, Vidaud M, Guibourdenche J, Luton D, Muller F, Bellet D, Giovagrandi Y, Tarrade A, Porquet D, Blot P, et al. Defect of villous cytotrophoblast differentiation into syncytiotrophoblast in Down's syndrome. The Journal of clinical endocrinology and metabolism 2000; 85:3700–7.

64. Pidoux G, Gerbaud P, Cocquebert M, Segond N, Badet J, Fournier T, Guibourdenche J, Evain-Brion D. Review: Human trophoblast fusion and differentiation: Lessons from trisomy 21 placenta. Placenta 2011; 33:S81–6.

65. Malassiné A, Frendo J-L, Evain-Brion D. Trisomy 21- affected placentas highlight prerequisite factors for human trophoblast fusion and differentiation. International Journal of Developmental Biology 2009; 54:475–82.

66. Bryant AH, Bevan RJ, Spencer-Harty S, Scott LM, Jones RH, Thornton CA. Expression and function of NOD-like receptors by human term gestation-associated tissues. Placenta 2017; 58:25–32.

67. Lappas M. NOD1 and NOD2 Regulate Proinflammatory and Prolabor Mediators in Human Fetal Membranes and Myometrium via Nuclear Factor-Kappa B. Biol Reprod 2013; 89:Article 14, 1-11.

68. Bourhis LL, Benko S, Girardin SE. Nod1 and Nod2 in innate immunity and human inflammatory disorders. Biochem Soc T 2007; 35:1479–84.

69. Marchand M, Horcajadas JA, Esteban FJ, McElroy SL, Fisher SJ, Giudice LC. Transcriptomic Signature of Trophoblast Differentiation in a Human Embryonic Stem Cell Model. Biology of Reproduction 2011; 84:1258–71.

70. Wang B, Palomares K, Parobchak N, Cece J, Rosen M, Nguyen A, Rosen T. Glucocorticoid Receptor Signaling Contributes to Constitutive Activation of the Noncanonical NF-κB Pathway in Term Human Placenta. Molecular Endocrinology 2013; 27:203–11.

71. Tang D, Kang R, Coyne CB, Zeh HJ, Lotze MT. PAMPs and DAMPs: signal 0s that spur autophagy and immunity. Immunological Reviews 2012; 249:158–75.

72. Rijn BB van, Franx A, Steegers EAP, Groot CJM de, Bertina RM, Pasterkamp G, Voorbij HAM, Bruinse HW, Roest M. Maternal TLR4 and NOD2 Gene Variants, Pro-Inflammatory Phenotype and Susceptibility to Early-Onset Preeclampsia and HELLP Syndrome. Plos One 2008; 3:e1865.

73. Robertson SA, Hutchinson MR, Rice KC, Chin P, Moldenhauer LM, Stark MJ, Olson DM, Keelan JA. Targeting Toll-like receptor-4 to tackle preterm birth and fetal inflammatory injury. Clin Transl Immunol 2020; 9:e1121.

74. Costello MJ, Joyce SK, Abrahams VM. NOD Protein Expression and Function in First Trimester Trophoblast Cells. Am J Reprod Immunol 2007; 57:67–80.

75. Strauss JF, Romero R, Gomez-Lopez N, Haymond-Thornburg H, Modi BP, Teves ME, Pearson LN, York TP, Schenkein HA. Spontaneous preterm birth: advances toward the discovery of genetic predisposition. Am J Obstet Gynecol 2018; 218:294-314.e2.

76. Härtel Ch, Finas D, Ahrens P, Kattner E, Schaible Th, Müller D, Segerer H, Albrecht K, Möller J, Diedrich K, et al. Polymorphisms of genes involved in innate immunity: association with preterm delivery. Mhr Basic Sci Reproductive Medicine 2004; 10:911–5.

77. Tekola-Ayele F, Zeng X, Ouidir M, Workalemahu T, Zhang C, Delahaye F, Wapner R. DNA methylation loci in placenta associated with birthweight and expression of genes relevant for early development and adult diseases. Clin Epigenetics 2020; 12:78.

78. Vennou KE, Kontou PI, Braliou GG, Bagos PG. Meta-analysis of gene expression profiles in preeclampsia. Pregnancy Hypertens 2020; 19:52–60.

79. Ramasammy R, Munisammy L, Sweta K, Selvakumar S, Velu K, Rani J, Kajalakshmy S. Association between GCK gene polymorphism and gestational diabetes mellitus and its pregnancy outcomes. Meta Gene 2021; 28:100856.

80. Beaumont RN, Warrington NM, Cavadino A, Tyrrell J, Nodzenski M, Horikoshi M, Geller F, Myhre R, Richmond RC, Paternoster L, et al. Genome wide associatin study of offspring birth weight in 86577 women identifies five novel loci and highlights maternal genetic effects that are independent of fetal genetics. Human Molecular Genetics 2018; 27:742–56.

81. Lang CT, Markham KB, Behrendt NJ, Suarez AA, Samuels P, Vandre DD, Robinson JM, Ackerman WE. Placental Dysferlin Expression is Reduced in Severe Preeclampsia. Placenta 2009; 30:711–8.

82. Varanou A, Withington SL, Lakasing L, Williamson C, Burton GJ, Hemberger M. The importance of cysteine cathepsin proteases for placental development. J Mol Med 2006; 84:305–17.

83. Jia R-Z, Zhang X, Hu P, Liu X-M, Hua X-D, Wang X, Ding H-J. Screening for differential methylation status in human placenta in preeclampsia using a CpG island plus promoter microarray. Int J Mol Med 2012; 30:133–41.

84. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, et al. Tissue-based map of the human proteome. Science 2015; 347:1260419.

85. Guo X, Zhang Y, Wang P, Li T, Fu W, Mo X, Shi T, Zhang Z, Chen Y, Ma D, et al. VSTM1-v2, a novel soluble glycoprotein, promotes the differentiation and activation of Th17 cells. Cell Immunol 2012; 278:136–42.

86. Sonderegger S, Pollheimer J, Knöfler M. Wnt Signalling in Implantation, Decidualisation and Placental Differentiation – Review. Placenta 2010; 31:839–47.

87. Knöfler M, Pollheimer J. Human placental trophoblast invasion and differentiation: a particular focus on Wnt signaling. Frontiers in Genetics 2013; 4:190.

88. Kokkinos MI, Murthi P, Wafai R, Thompson EW, Newgreen DF. Cadherins in the human placenta – epithelial–mesenchymal transition (EMT) and placental development. Placenta 2010; 31:747–55.

89. Adu-Gyamfi EA, Czika A, Gorleku PN, Ullah A, Panhwar Z, Ruan L-L, Ding Y-B, Wang Y-X. The Involvement of Cell Adhesion Molecules, Tight Junctions, and Gap Junctions in Human Placentation. Reprod Sci 2021; 28:305–20.

90. Yeung KR, Chiu CL, Pidsley R, Makris A, Hennessy A, Lind JM. DNA methylation profiles in preeclampsia and healthy control placentas. Am J Physiol-heart C 2016; 310:H1295–303.

91. Raj K, Horvath S. Current perspectives on the cellular and molecular features of epigenetic ageing. Exp Biol Med 2020; 245:1532–42.

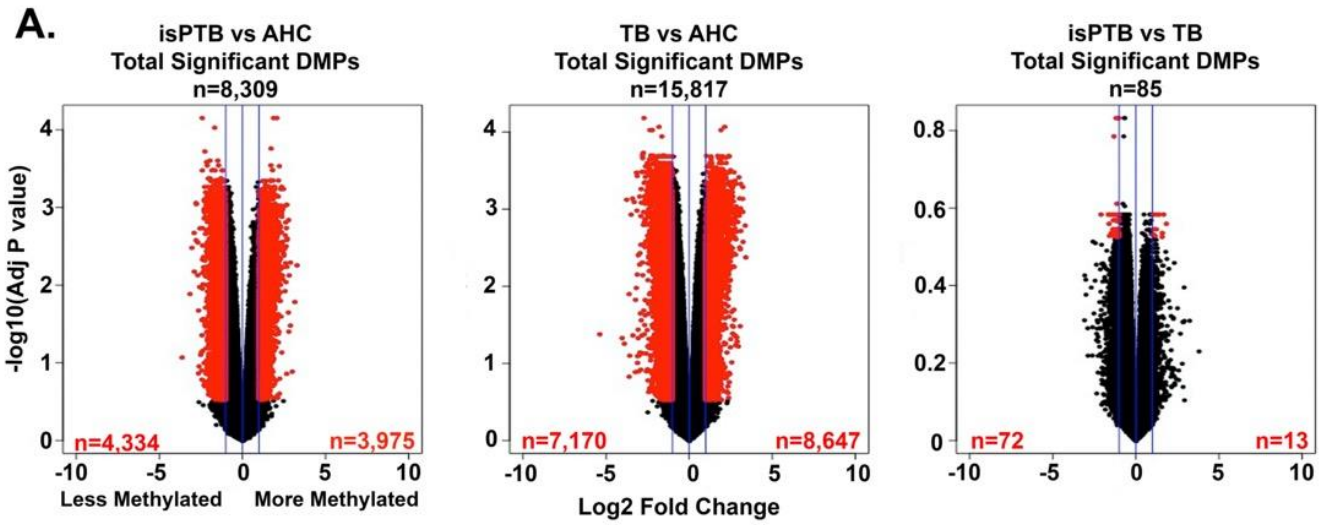
92. Lee Y, Choufani S, Weksberg R, Wilson SL, Yuan V, Burt A, Marsit C, Lu AT, Ritz B, Bohlin J, et al. Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels. Aging 2019; 11:4238–53.

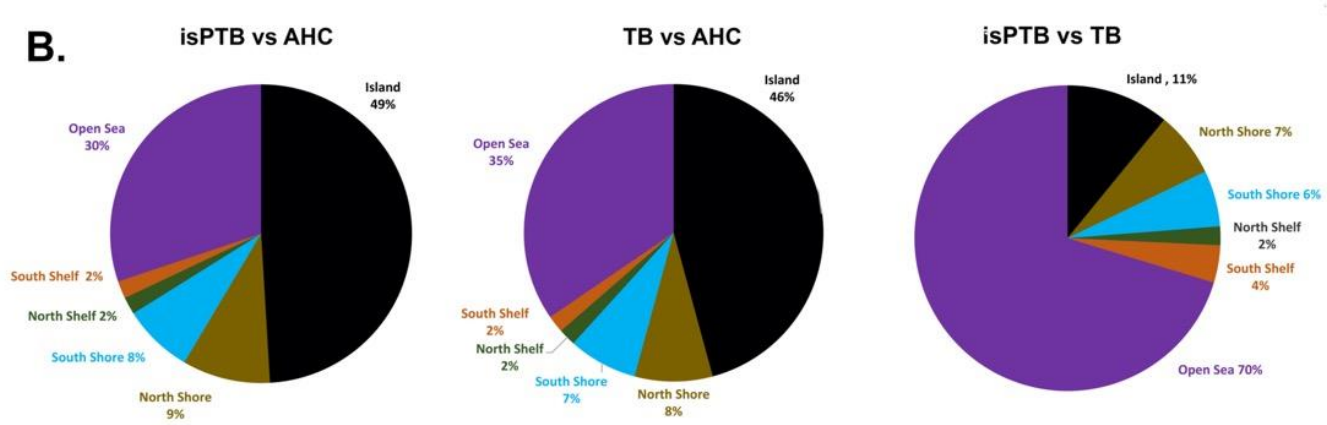
Figures and Tables:

Table 1: Clinical characteristics of the placental villous samples included in the methylation analyses

Characteristics	Acute Histological Chorioamnionitis Births (AHC)	Idiopathic Spontaneous Preterm Births (isPTB)	Term Births	P-values
Number of samples	8	11	8	
Maternal Age	34.5(25-40)	25(18-39)	28(19-37)	NS[1]
Gestational Age	32(29-35)*	33(30-36)*	39(38-41)	<0.0001[1]
Fetal sex (% female)	3(38%)	6(55%)	4(38%)	NS[2]
Fetal weight (grams)	1765(1360-2300)*	2105(1450-2722)*	3820(3650-4527)	<0.0001[1]
Birth weight percentile	55(20-80)	60(3-80)	90(60-99)	NS[1]
SGA %	0	18.0%	0	
Delivery type				
Cesarean (%)	4(50%)	4(37%)	5(50%)	NS[2]
Infection Status				
(% Positive)	8(100%)*	0(0%)	0(0%)	<0.0001[2]

Data shown as median with range or total number with percent
[1]ANOVA with Tukey's correction for multiple comparisons
[2]Chi Square Analyses
NS=Not significant





Open sea = >4kb away from CpG island/isolated CpG sites
Island = within the CpG Island
Shelf = 2-4kb flankir
Shore = 0-2kb flankir

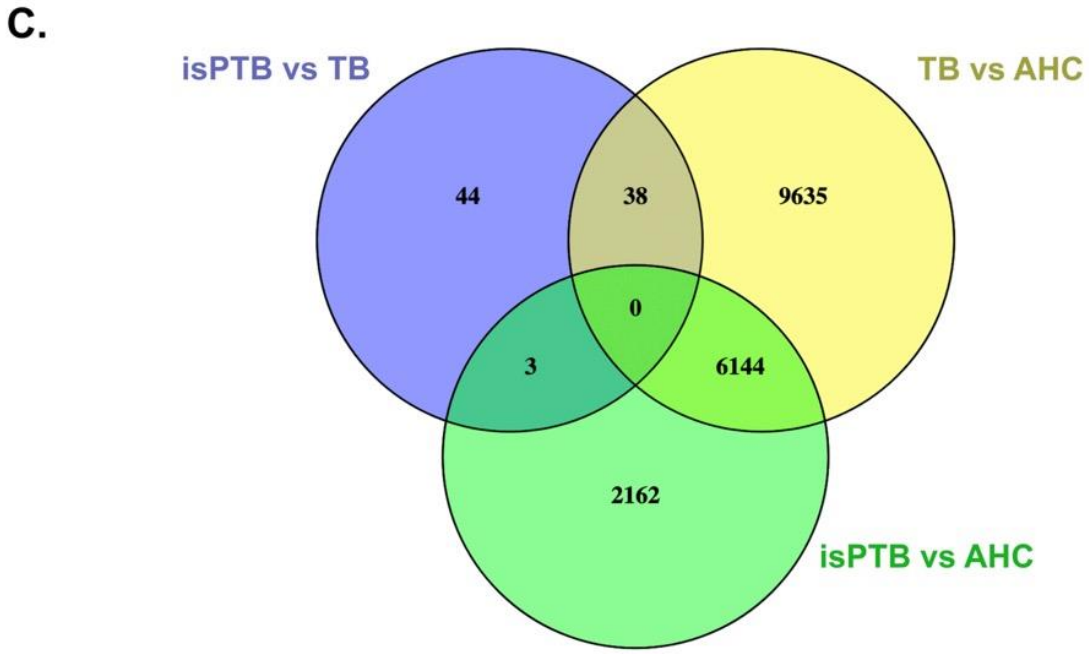


Figure 1: Identification of methylation signatures using a comparative approach. A. Differentially methylated positions were identified using pairwise comparisons in *limma*. Red points indicate significant DMPs with a threshold of log2 fold change >1 and Benjamini Hochberg adjusted p value <0.2. Blue lines represent log2fold change of 1. **B.** Genomic distribution of DMPs in the pairwise comparisons. The majority of DMPs in the isPTB and TB verses AHC comparisons are located inside or close to known CpG islands. However, in the isPTB verses TB comparison, the majority of DMPs are in open sea regions with no known islands within 4kb. **C.** The venn diagram represents the intersection of pairwise comparisons to classify significant DMPs into isPTB and AHC specific methylation signatures.

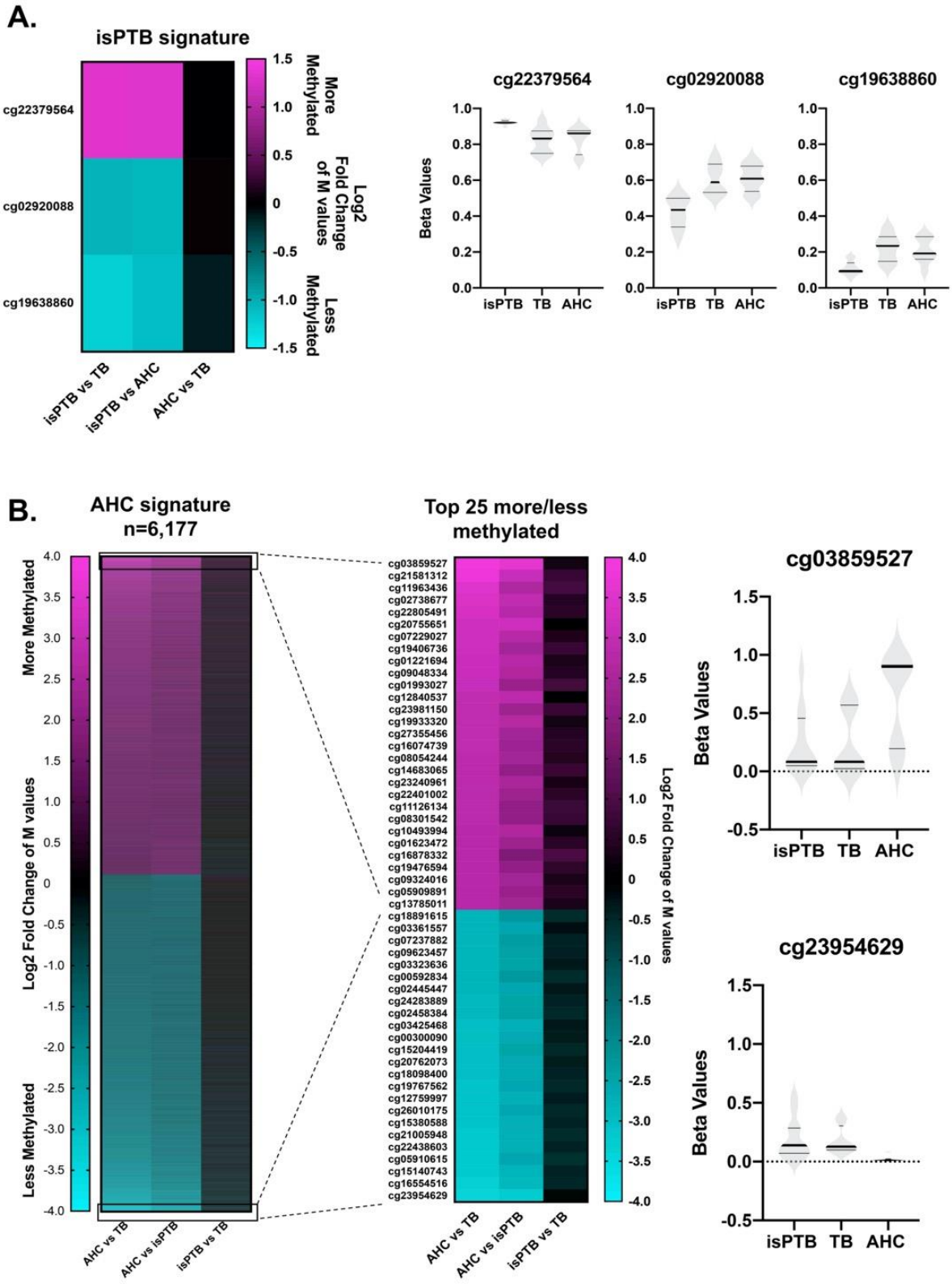


Figure 2: Identification of significant methylation signatures for isPTB and AHC DMPs.
A. Three DMPs identified as having a isPTB specific methylation pattern where the isPTB samples were more or less methylated compared to the AHC or TB samples. **B.** 6,177 DMPs demonstrating a methylation pattern where the AHC samples were more or less methylated than the isPTB or TB samples. The breakout heatmap shows the pattern or the top 25 more and less methylated samples and demonstrates the similarity of methylation between the isPTB and TB samples. The distribution of individual sample beta values was assessed to determine if there were outliers or artifacts influencing the methylation patterns

Table 2 Summary of significantly differentiated DMRs identified by *DMRcate* encompassing both coding and non-coding loci

Pairwise comparison	*Number of Significant DMRs Identified**	*Width of DMR (Range)*	*Number of Significant Probes in DMR (Range)*
isPTB vs TB	56	180-1750bp	5-18 probes
isPTB vs AHC	12,883	83-9,386bp	5-110 probes
TB vs AHC	19,006	37-14,383bp	5-202 probes

*minimum smoothed FDR <0.05

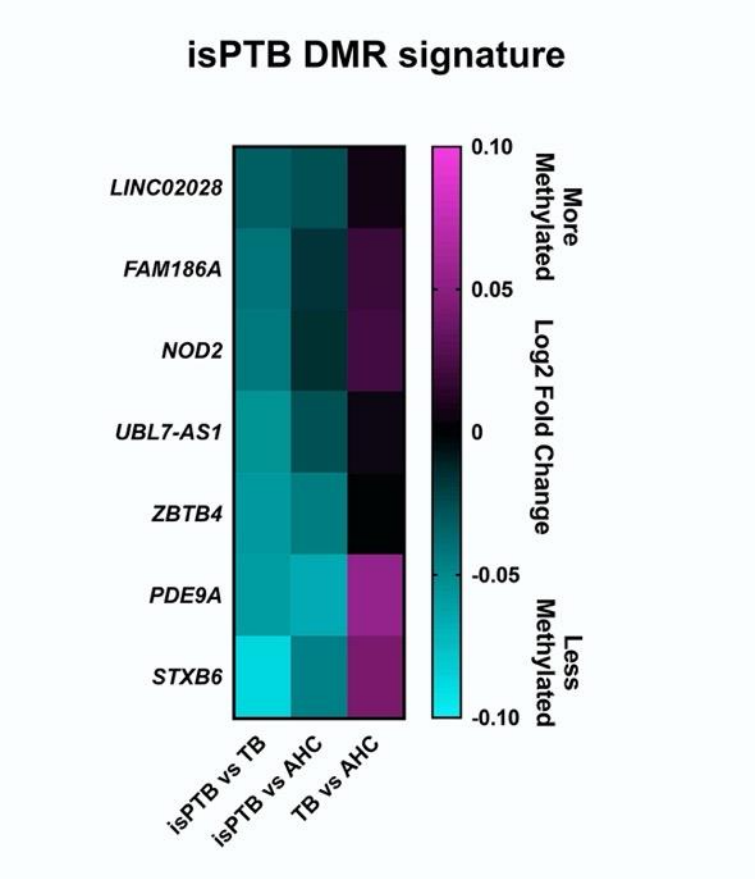


Figure 4 isPTB specific DMR signature. Differentially methylated DMRs were identified by differences in the mean of the probe values across the DMR. Only 7 isPTB DMRs had an isPTB specific signature where the isPTB DMRs were less methylated than the TB or AHC DMRs. Two of the DMRs overlap non-coding regions. No DMRs were identified that were more methylated.

Table 3 Summary of isPTB signature DMRs

	Mean Difference of all probes in DMR			
Locus	**isPTB vs TB**	**isPTB vs AHC**	**TB vs AHC**	**DMR coordinates**
LINC02028	-0.033	-0.028	0.007	chr3:194072066-194072416
FAM186A	-0.0416	-0.0175	0.0192	chr12:50343856-50344626
NOD2	-0.043	-0.015	0.022	chr16:50715192-50715700
UBL7-AS1	-0.054	-0.028	0.005	chr15:74466794-74467158
ZBTB4	-0.058	-0.045	0.0008	chr17:7461421-7462028
PDE9A	-0.059	-0.066	0.054	chr21:42733397-42733894
STXB6	-0.087	-0.0466	0.042	chr14:24808650-24810213

Table 4 Functional information for the isPTB DMRs

Locus	Overlaps with CpG Island	Location
LINC02028	chr3:194070715-194071468	Promoter
FAM186A	NA	Intronic
NOD2	NA	Intronic
UBL7-AS1	NA	Intronic
ZBTB4	NA	3'UTR/last exon
PDE9A	NA	Intron/exon boundary
STXB6	NA	3'UTR

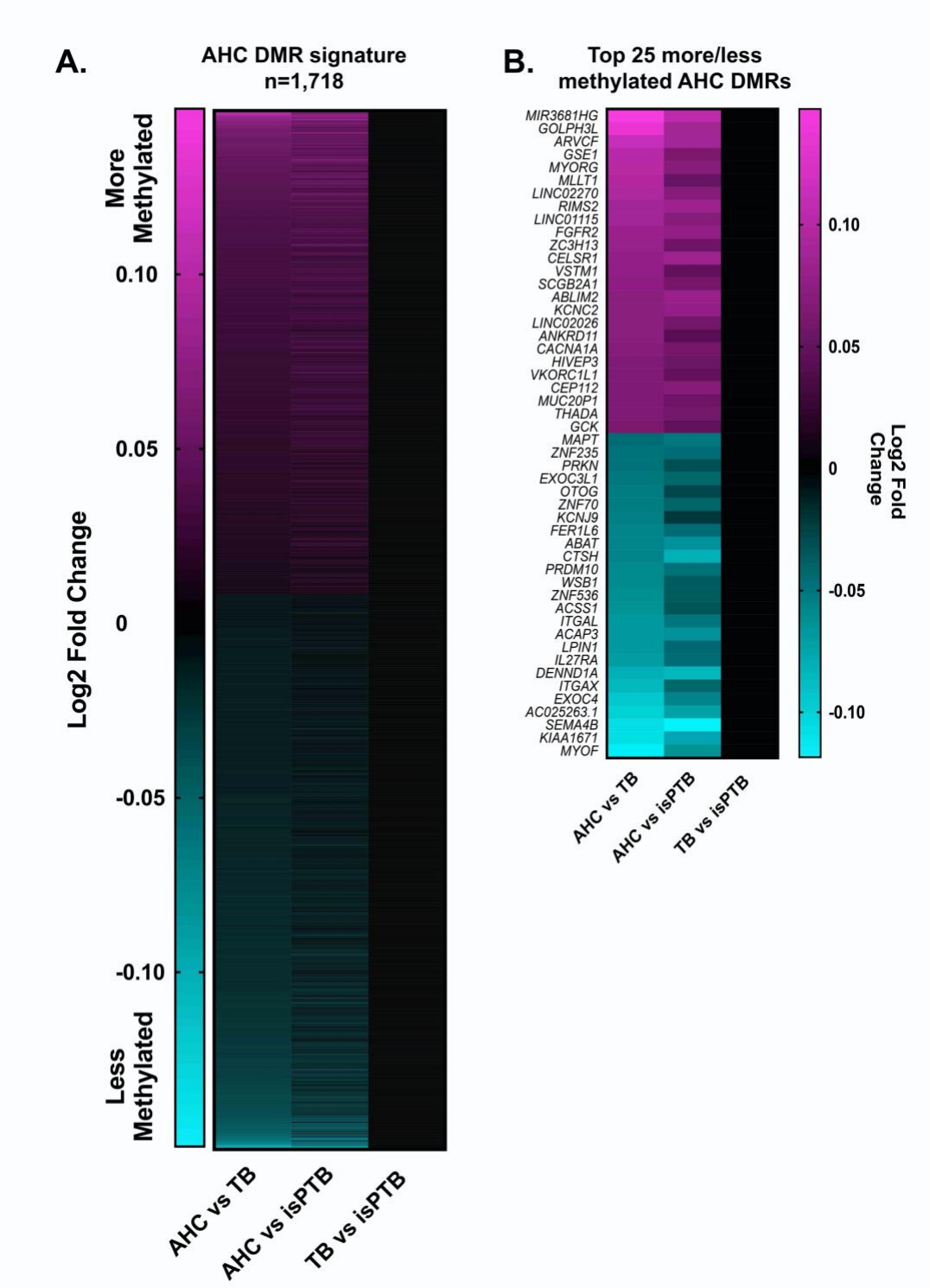


Figure 5 AHC specific DMR signature A. Differentially methylated DMRs were identified by differences in the mean of the probe values across the DMR. AHC specific DMRs are defined by when the AHC DMRs were More or less methylated than the TB or isPTB DMRs. **B.** The top 25 more and less methylated DMRs demonstrates the clarity of the molecular signature, as there is no significant differential methylation in the TB vs isPTB comparison.

Table 5: The top 25 more and less methylated DMR mean differences in pairwise comparisons

DMR location	*Locus Name*	*Mean Diff AHC vs TB*	*Mean Diff AHC v isPTB*	*Mean Diff TB v isPTB*
chr2:11915711-11916260	***MIR3681HG***	0.1474	0.1065	Not significant
chr1:150692971-150694343	***GOLPH3L***	0.1366	0.0874	Not significant
chr22:19973978-19975691	***ARVCF***	0.1125	0.0887	Not significant
chr16:85342729-85343936	***GSE1***	0.1023	0.0615	Not significant
chr9:34372089-34373067	***MYORG***	0.0993	0.0698	Not significant
chr19:6230050-6230665	***MLLT1***	0.0986	0.0519	Not significant
chr4:12224743-12225077	***LINC02270***	0.0946	0.0679	Not significant
chr8:103750821-103751623	***RIMS2***	0.0894	0.0832	Not significant
chr2:794646-796536	***LINC01115***	0.0881	0.0683	Not significant
chr10:121577971-121579007	***FGFR2***	0.0812	0.0774	Not significant
chr13:45965025-45966279	***ZC3H13***	0.0800	0.0565	Not significant
chr22:46440394-46442103	***CELSR1***	0.0786	0.0848	Not significant
chr19:54040774-54041856	***VSTM1***	0.0753	0.0491	Not significant
chr11:62211493-62212431	***SCGB2A1***	0.0748	0.0606	Not significant
chr4:7967275-7969643	***ABLIM2***	0.0741	0.0827	Not significant
chr12:75057893-75058468	***KCNC2***	0.0723	0.0780	Not significant
chr12:126018024-126018364	***AC005186.1***	0.0721	0.0576	Not significant
chr16:89488412-89489377	***ANKRD11***	0.0714	0.0426	Not significant
chr19:13616871-13617970	***CACNA1A***	0.0678	0.0609	Not significant
chr1:41831580-41832649	***HIVEP3***	0.0668	0.0545	Not significant
chr7:65878352-65879115	***VKORC1L1***	0.0667	0.0491	Not significant
chr17:66097276-66098113	***CEP112***	0.0665	0.0687	Not significant
chr3:195619562-195620147	***MUC20P1***	0.0653	0.0550	Not significant
chr2:43327937-43328914	***THADA***	0.0647	0.0588	Not significant
chr7:44152238-44154322	***GCK***	0.0632	0.0475	Not significant
chr17:46018654-46019184	***MAPT***	-0.0470	-0.0518	Not significant
chr19:44302666-44303858	***ZNF235***	-0.0494	-0.0474	Not significant
chr6:161560605-161561121	***PRKN***	-0.0494	-0.0327	Not significant

chr16:67184164-67185527	***EXOC3L1***	-0.0506	-0.0433	Not significant
chr11:17568197-17569556	***OTOG***	-0.0522	-0.0294	Not significant
chr22:23744094-23745131	***ZNF70***	-0.0536	-0.0400	Not significant
chr1:160084263-160085568	***KCNJ9***	-0.0541	-0.0222	Not significant
chr8:123859056-123859953	***FER1L6***	-0.0563	-0.0473	Not significant
chr16:8724073-8724983	***ABAT***	-0.0577	-0.0662	Not significant
chr15:78933106-78934580	***CTSH***	-0.0579	-0.0794	Not significant
chr11:129993525-129993935	***PRDM10***	-0.0596	-0.0476	Not significant
chr17:27312855-27313499	***WSB1***	-0.0598	-0.0373	Not significant
chr19:30413468-30414886	***ZNF536***	-0.0621	-0.0370	Not significant
chr20:25013229-25014771	***ACSS1***	-0.0649	-0.0353	Not significant
chr16:30485296-30485966	***ITGAL***	-0.0683	-0.0504	Not significant
chr1:1296671-1297807	***ACAP3***	-0.0685	-0.0662	Not significant
chr2:11679584-11680144	***LPIN1***	-0.0691	-0.0437	Not significant
chr19:14048977-14049823	***IL27RA***	-0.0702	-0.0460	Not significant
chr9:123656764-123657427	***DENND1A***	-0.0794	-0.0852	Not significant
chr16:31366142-31366536	***ITGAX***	-0.0852	-0.0428	Not significant
chr7:133811022-133812369	***EXOC4***	-0.0945	-0.0578	Not significant
chr12:69724920-69725444	***AC025263.1***	-0.0988	-0.0738	Not significant
chr15:90208739-90209326	***SEMA4B***	-0.1083	-0.1185	Not significant
chr22:24988020-24990749	***KIAA1671***	-0.1093	-0.0773	Not significant
chr10:93334974-93335677	***MYOF***	-0.1173	-0.0643	Not significant

Table 6: Bioinformatic functional assessment of more methylated AHC signature DMRs via PantherDB

	Homo sapiens (all genes in database)	*Genes from input list*	*Expected*	*Fold Enrichment*	*Adjusted P-Value**
PANTHER Pathways					
Cadherin signaling pathway (P00012)	164	21	5.34	3.94	6.51E-05
Wnt signaling pathway (P00057)	317	30	10.31	2.91	1.03E-04
GO biological process complete					
homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156)	168	26	5.47	4.76	4.62E-06
cell-cell adhesion via plasma-membrane adhesion molecules (GO:0098742)	257	28	8.36	3.35	1.05E-03
GO molecular function complete					
ion binding (GO:0043167)	6354	277	206.71	1.34	5.61E-05
binding (GO:0005488)	16539	593	538.05	1.1	8.90E-05
molecular_function (GO:0003674)	18245	631	593.55	1.06	4.23E-03
metal ion binding (GO:0046872)	4268	192	138.85	1.38	4.82E-03
cation binding (GO:0043169)	4354	194	141.65	1.37	9.08E-03
adenyl nucleotide binding (GO:0030554)	1572	84	51.14	1.64	3.90E-02

*Fisher Test Bonferroni Corrected for multiple comparisons

Table 7: Bioinformatic functional assessment of less methylated AHC signature DMRs via PantherDB

	Homo sapiens (all genes in database)	*Genes from input list*	*Expected*	*Fold Enrichment*	*Adjusted P-Value**
GO biological process complete					
cell morphogenesis involved in differentiation (GO:0000904)	568	49	21.68	2.26	5.15E-03
detection of chemical stimulus (GO:0009593)	522	2	19.92	0.1	8.02E-03
cell morphogenesis (GO:0000902)	721	56	27.52	2.04	1.96E-02
detection of chemical stimulus involved in sensory perception (GO:0050907)	486	2	18.55	0.11	3.64E-02
GO molecular function complete					
binding (GO:0005488)	16539	689	631.2	1.09	2.56E-04
protein binding (GO:0005515)	14359	615	548.01	1.12	4.39E-04
molecular_function (GO:0003674)	18245	739	696.31	1.06	1.33E-03
ion binding (GO:0043167)	6354	310	242.5	1.28	1.69E-03
olfactory receptor activity (GO:0004984)	441	2	16.83	0.12	4.87E-02

*Fischer Test Bonferroni Corrected for multiple comparison

SUPPLEMENTAL TABLES (some will be in Excel)

Supplemental Table 1 Probe Filtering during quality control assessment

Total probes read into pipeline		**866,901**
Failed Detection	810	866,091
Normalization (FunNorm)	232	865,859
Filtering probes after quality control		
Probes that failed in 2+ samples	8,546	**865,859**
Remove X/Y probes	18,913	857,313
Remove SNPs (Manifest)	28,488	838,400
Remove SNPs (Zhou 2016)	13,303	809,912
Remove Cross hybridizing probes (McCartney 2016)	38,280	796,609
Remove Blacklist probes (2019 Blacklist)	119	758,329
Total probes left for analyses		**758,210**

Supplemental Table 2 Statistical testing in *limma* to determine significant DMPs between pairwise comparisons

	BH adjusted p <0.05			**BH adjusted p <0.1**		
	isPTB vs AHC	**TB vs AHC**	**isPTB vs TB**	**isPTB vs AHC**	**TB vs AHC**	**isPTB vs TB**
More methylated probes	13,111	41,767	0	27,797	71,566	0
Less methylated probes	17,037	31,632	0	36,791	59,535	0
Total DMPs	30,148	73,399	0	64,588	131,101	0

	BH adjusted p <0.2			**BH adjusted p <0.3**		
	isPTB vs AHC	**TB vs AHC**	**isPTB vs TB**	**isPTB vs AHC**	**TB vs AHC**	**isPTB vs TB**
More methylated probes	51,382	116,150	0	73,338	152,935	29
Less methylated probes	72,136	109,987	7	105,617	152,647	593
Total DMPs	123,518	226,137	7	178,955	305,582	662

*Separate was selected within limma as the statistical method within limma
** Limma only selected for adjusted p value, not log2 fold change.

Supplemental Table 3 Values for isPTB DMPs (EXCEL)- this is a large file
Supplemental Table 4 Values for AHC DMP (EXCEL)- this is a large file

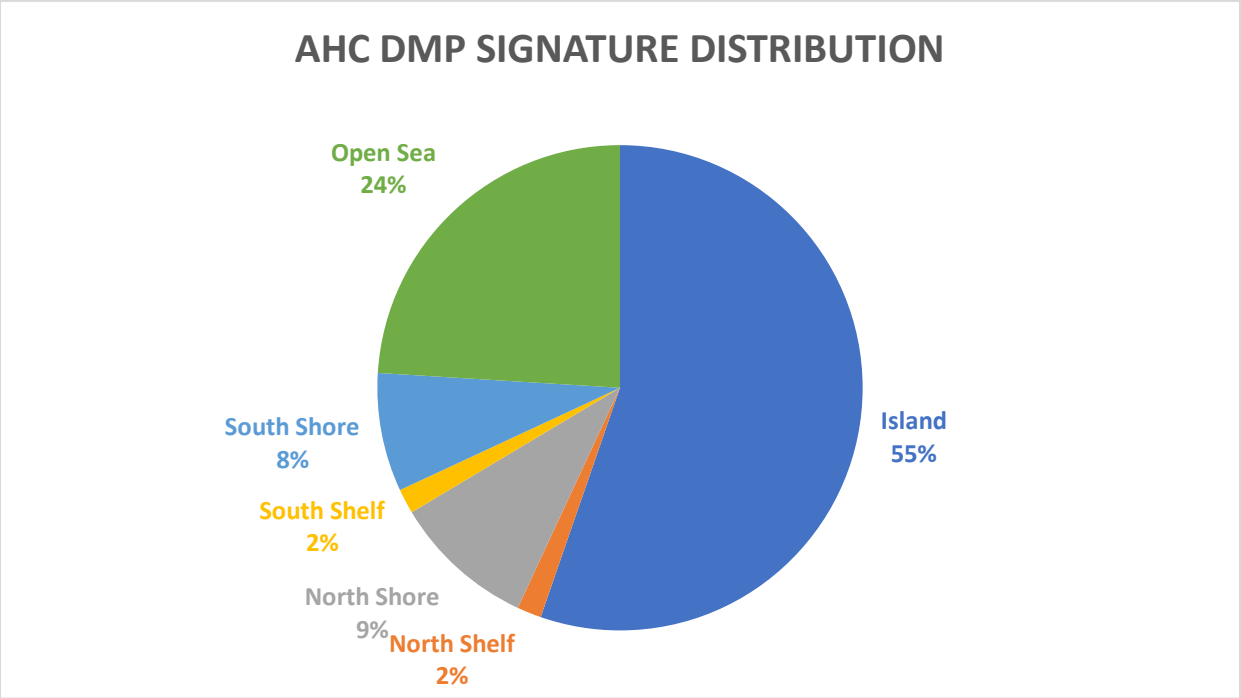
Supplemental Table 5 Statistical testing in *DMRcate* to determine significant DMPs between pairwise comparisons

BH adjusted p	<0.05	<0.2	<0.3	<0.5
isPTB vs TB	0	7	662	14,611
isPTB vs AHC	30,148	123,518	178,955	300,625
TB vs AHC	73,399	226,137	305,582	483,450

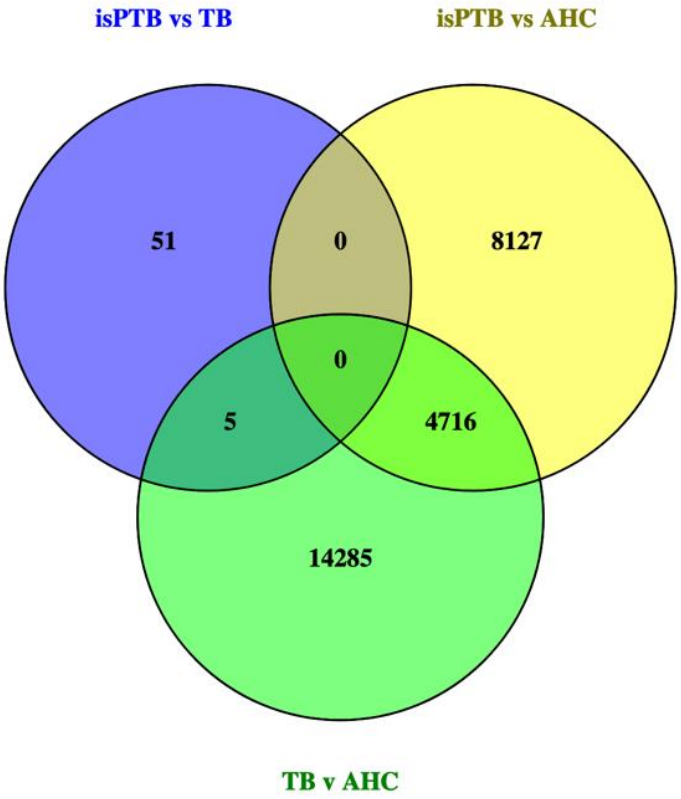
** Limma only selected for adjusted p value, not log2 fold change.

Supplemental Table 6 Values for AHC DMRs (EXCEL)- this is a large file

SUPPLEMENTAL FIGURES



Supplemental Figure 1: Genomic Distribution of DMPs within the AHC methylation signature. The distribution of 6,177 DMPs in the AHC signature DMPs. The majority of probes are found within CpG islands or closely associated with islands.



Supplemental Figure 2: Intersection of significant DMRs The venn diagram representing the intersection of pairwise comparisons to classify significant DMRs into isPTB and AHC specific signatures