Clinical Epigenetics

RESEARCH   **Open Access**

CrossMark

Analysis of two birth tissues provides new insights into the epigenetic landscape of neonates born preterm

Yonghui Wu[1†], Xinyi Lin[1,2†], Ives Yubin Lim[1†], Li Chen[1], Ai Ling Teh[1], Julia L. MacIsaac[3], Kok Hian Tan[4], Michael S. Kobor[3], Yap Seng Chong[1,5], Peter D. Gluckman[1,6] and Neerja Karnani[1,7*]

Abstract

Background: Preterm birth (PTB), defined as child birth before completion of 37 weeks of gestation, is a major challenge in perinatal health care and can bear long-term medical and financial burden. Over a million children die each year due to PTB complications, and those who survive can face developmental delays. Unfortunately, our understanding of the molecular pathways associated with PTB remains limited. There is a growing body of evidence suggesting the role of DNA methylation (DNAm) in mediating the effects of PTB on future health outcomes. Thus, epigenome-wide association studies (EWAS), where DNAm sites are examined for associations with PTB, can help shed light on the biological mechanisms linking the two.

Results: In an Asian cohort of 1019 infants (68 preterm, 951 full term), we examined and compared the associations between PTB and genome-wide DNAm profiles using both cord tissue ($n = 1019$) and cord blood ($n = 332$) samples on Infinium HumanMethylation450 arrays. PTB was significantly associated ($P < 5.8e{-}7$) with DNAm at 296 CpGs (209 genes) in the cord blood. Over 95% of these CpGs were replicated in other PTB/gestational age EWAS conducted in (cord) blood. This replication was apparent even across populations of different ethnic origin (Asians, Caucasians, and African Americans). More than a third of these 296 CpGs were replicated in at least 4 independent studies, thereby identifying a robust set of PTB-linked epigenetic signatures in cord blood. Interrogation of cord tissue in addition to cord blood provided novel insights into the epigenetic status of the neonates born preterm. Overall, 994 CpGs (608 genes, $P < 3.7e{-}7$) associated with PTB in cord tissue, of which only 10 of these CpGs were identified in the analysis using cord blood. Genes from cord tissue showed enrichment of molecular pathways related to fetal growth and development, while those from cord blood showed enrichment of immune response pathways. A substantial number of PTB-associated CpGs from both the birth tissues were also associated with gestational age.

Conclusions: Our findings provide insights into the epigenetic landscape of neonates born preterm, and that its status is captured more comprehensively by interrogation of more than one neonatal tissue in tandem. Both these neonatal tissues are clinically relevant in their unique ways and require careful consideration in identification of biomarkers related to PTB and gestational age.

Trial registration: This birth cohort is a prospective observational study designed to study the developmental origins of health and disease, and was retrospectively registered on 1 July 2010 under the identifier NCT01174875.

Keywords: Epigenome wide association study, Preterm birth, Gestational age, Tissue specificity, DNA methylation, Neonate

* Correspondence: neerja_karnani@sics.a-star.edu.sg
†Yonghui Wu, Xinyi Lin and Ives Yubin Lim contributed equally to this work.
¹Singapore Institute for Clinical Sciences, A*STAR, 30 Medical Drive,
Singapore 117609, Singapore
⁷Department of Biochemistry, Yong Loo Lin School of Medicine, National
University of Singapore, Singapore, Singapore
Full list of author information is available at the end of the article

Wu *et al. Clinical Epigenetics*   (2019) 11:26

Page 2 of 12

Background

Preterm birth (PTB), defined as delivery of the offspring before completion of 37 weeks of gestation, is a major public health problem that exerts a significant disease burden globally [1]. In 2016, World Health Organization estimated 15 million babies (at least 1 in 10 babies) to be born preterm annually, and that these numbers are rising each year [2]. PTB is associated with developmental delays, and infants born preterm are at an increased risk of mortality from infancy to adulthood due to the onset of various chronic health problems [3, 4]. However, the biological pathways underlying the associations between PTB and future health remain elusive [5, 6]. Epigenetic mechanisms play a critical role in regulating cell lineage commitment and fetal programing and are highly sensitive to in utero perturbations. Any interference with the epigenetic settings within the cell or its developmental state can have life-long impact on the health of the offspring. Thus, epigenome-wide association studies (EWAS) related to PTB [7–12] can help elucidate the biological mechanisms linking the two [13].

There is a growing body of evidence suggesting the influence of PTB on neonatal epigenome through DNA methylation (DNAm) [7, 8, 13–18]. Earlier efforts in interrogating DNAm changes in association with PTB typically focused on candidate regions of the epigenome [19, 20] or were conducted in smaller sample sizes [7–10, 18]. Recently, some research groups have conducted EWAS of gestational age (GA), with some using larger sample sizes. Schroeder et al. [21] reported and replicated the association between DNAm and GA at CpG sites in 25 genes, genes previously implicated in labor and delivery and adverse health outcomes. Lee et al. [22] reported DNAm at three regions associated with GA, regions located near genes that play key roles in fetal development (*NFIX*, *RAPGEF2*, *MSRB3*). Bohlin et al. [11] and Simpkin et al. [12] reported DNAm at 5474 CpG sites and 224 CpG sites to associate with GA, respectively. Though the total sample sizes in these GA EWAS were larger, with the exception of Bohlin et al. [11], the number of preterm infants in the analyses did not exceed 30.

While earlier studies have made significant progress in identifying DNAm perturbations associated with PTB/GA and enhanced our understanding of the epigenetic processes associated with PTB, a few important considerations remain. First, as earlier investigations were primarily conducted in Caucasian and/or African American populations, it is unclear how these findings hold in an Asian population. Second, earlier work primarily focused on examination of DNAm in infant cord blood [7, 9–12, 16, 21, 22], but there have been no studies done on cord tissue. Since cord tissue and cord blood originate from different cell lineages, each tissue potentially reveals unique perspectives within the preterm scenario. Pertinently, our earlier work has demonstrated that neonate EWAS conducted using infant cord tissue can give very distinct findings from those conducted in cord blood [23]. Hence, the two tissues together capture a better understanding of the epigenetic alterations induced by a suboptimal fetal environment. Here, we present the first EWAS of PTB conducted in an Asian cohort, where we examine and compare the associations of PTB with DNAm in both infant cord tissue and cord blood.
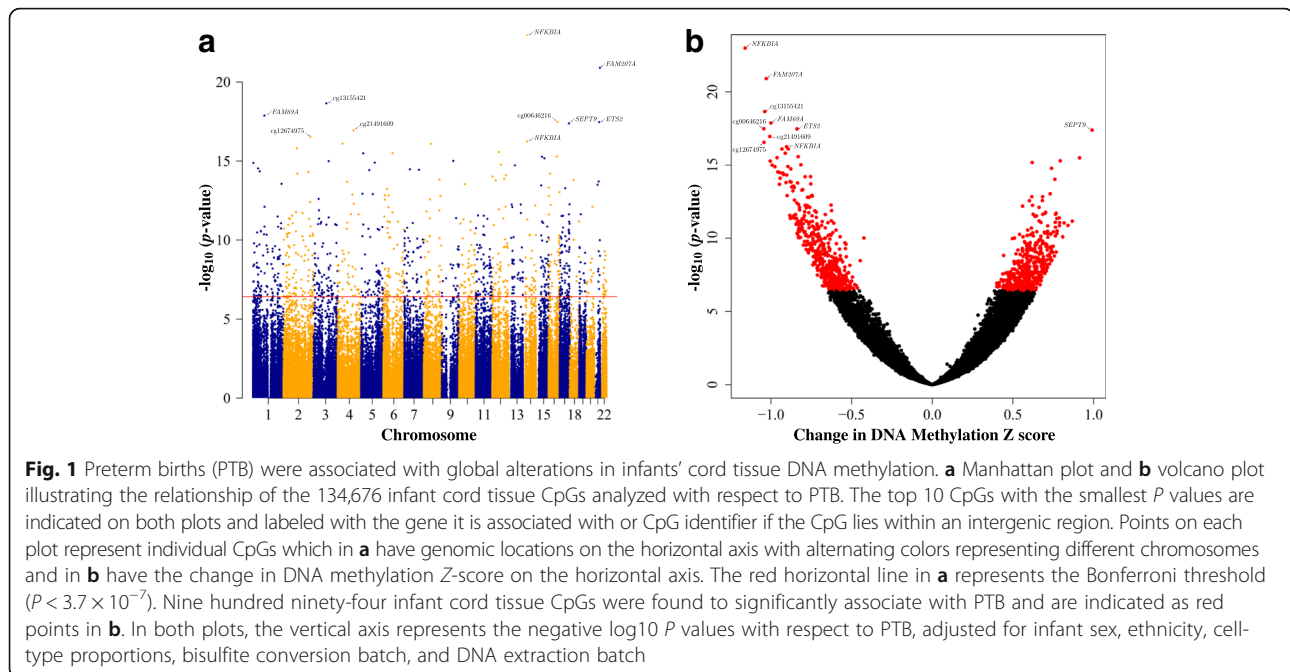
Results

Study population

This study involved 1019 infants from live singleton births, of which 68 infants were born preterm (Additional file 1: Figure S1A). Summary statistics of these infants are provided in Additional file 2: Table S1. The ethnic distribution of study subjects with available cord tissue samples was 58% Chinese, 25% Malay, and 17% Indian. Fifty-three percent of the infants were male. The difference in the distributions of ethnicity ($P = 0.88$) and sex ($P = 0.90$) of the infants in preterm vs. term groups was not statistically significant. We interrogated DNAm profiles derived from infant cord tissue and cord blood using the Infinium Human-Methylation450 array. DNAm data was available for all 1019 infants for cord tissue and in a subset of infants for cord blood (332 infants, including 31 preterm infants, Additional file 2: Table S2, Additional file 1: Figure S1B). Similarly, the distributions of infants with cord blood samples in preterm vs. term groups were not significantly different with respect to ethnicity ($P = 0.47$) and infant sex ($P = 0.58$). After quality control and elimination of CpGs with low variability, 134,676 and 85,624 CpGs were retained for subsequent analyses in cord tissue and cord blood, respectively.

Cord tissue reflected extensive associations between PTB and infant DNAm

We examined the association between cord tissue DNAm and PTB and identified 994 CpGs to be significantly associated with PTB using a Bonferroni multiple testing correction ($P < 3.7\mathrm{e}{-7}$; Fig. 1, Additional file 2: Table S3). The percentage of PTB-associated CpGs in hypomethylation (49%, 492 CpGs) and hypermethylation (51%, 502 CpGs) groups was almost equal (Fig. 1b), and their absolute effect size estimates (change in cord tissue DNAm Z-score with respect to PTB status) ranged from 0.40 to 1.16 (Additional file 2: Table S3). These 994 CpGs mapped to 608 unique genes, with the top most statistically significant CpGs mapping to several transcription factors such as nuclear factor of kappa light polypeptide gene enhancer in B cells

Fig. 1 Preterm births (PTB) were associated with global alterations in infants' cord tissue DNA methylation. **a** Manhattan plot and **b** volcano plot illustrating the relationship of the 134,676 infant cord tissue CpGs analyzed with respect to PTB. The top 10 CpGs with the smallest *P* values are indicated on both plots and labeled with the gene it is associated with or CpG identifier if the CpG lies within an intergenic region. Points on each plot represent individual CpGs which in **a** have genomic locations on the horizontal axis with alternating colors representing different chromosomes and in **b** have the change in DNA methylation *Z*-score on the horizontal axis. The red horizontal line in **a** represents the Bonferroni threshold ($P < 3.7 \times 10^{-7}$). Nine hundred ninety-four infant cord tissue CpGs were found to significantly associate with PTB and are indicated as red points in **b**. In both plots, the vertical axis represents the negative log10 *P* values with respect to PTB, adjusted for infant sex, ethnicity, cell-type proportions, bisulfite conversion batch, and DNA extraction batch

inhibitor, alpha (*NFKBIA*); ETS proto-oncogene 2, transcription factor (*ETS2*); and potential cell cycle control factors such as Septin 9 (*SEPT9*), family with sequence similarity 69 member A (*FAM69A*), and sequence similarity 207 member A (*FAM207A*). These 994 cord tissue CpGs remained largely statistically significant in sensitivity analyses (Additional file 1: Figure S2, Additional file 2: Table S3), with 79–87% remaining statistically significant after Bonferroni adjustment and 99–100% reflecting nominal significance with *P* value $< 10^{-4}$.
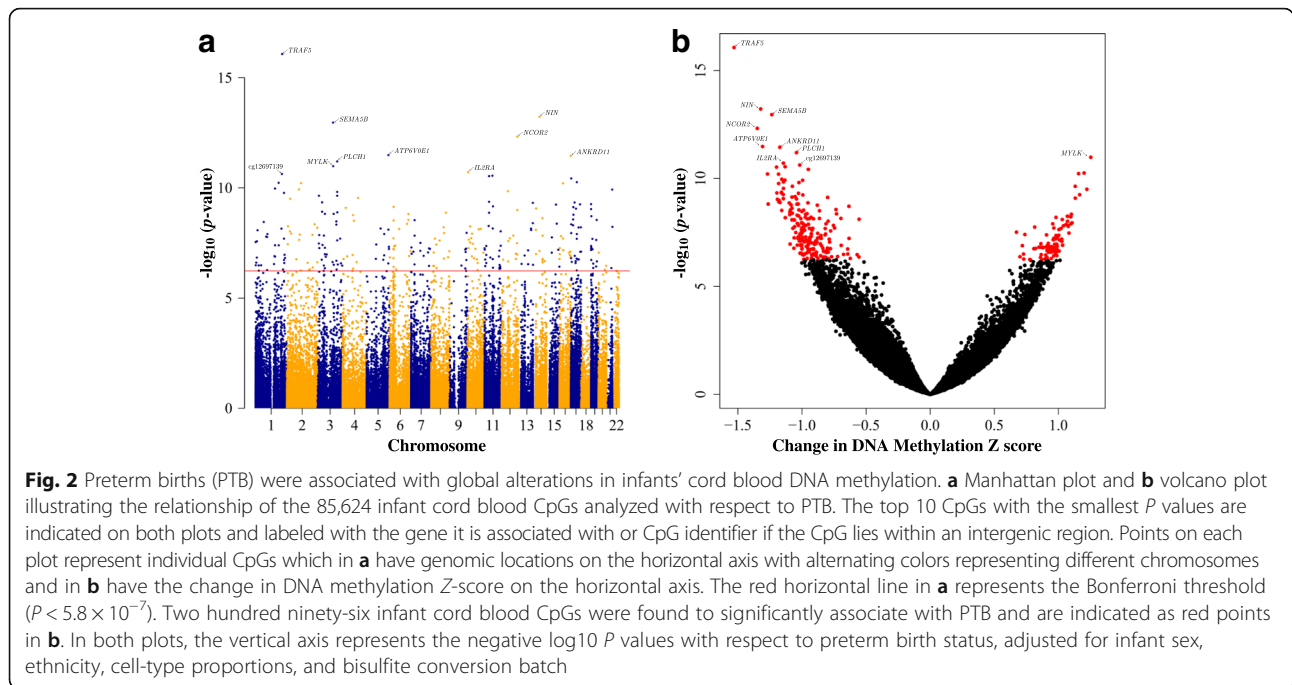
Cord blood reflected extensive associations between PTB and infant DNAm

We further interrogated the association between cord blood DNAm and PTB in 332 infants (these 332 infants are a subset of the 1019 infants). After adjusting for multiple testing using a Bonferroni correction ($P < 5.8e-7$; Fig. 2, Additional file 2: Table S4), 296 CpGs in 209 unique genes were identified to significantly associate with PTB in infant cord blood. These CpGs had absolute effect size estimates (change in cord blood DNAm *Z*-score with respect to PTB status) ranging from 0.55 to 1.53 (Additional file 2: Table S4). Ten CpGs overlapped between the cord blood (296 CpGs) and cord tissue (994 CpGs) analyses. The top most statistically significant CpGs on this list included immune response and signaling genes such as TNF receptor-associated factor 5 (*TRAF5*), nuclear receptor corepressor 2 (*NCOR2*), myosin light-chain kinase (*MYLK*), and interleukin 2 receptor subunit alpha (*IL2RA*) and phospholipase C eta 1 (*PLCH1*). These 296 cord blood CpGs remained largely statistically significant in sensitivity analyses (Additional file 1:

Figure S3, Additional file 2: Table S4), with 94–97% remaining Bonferroni significant and 100% reflecting nominal significance with *P* value $< 10^{-4}$. In contrast to our observation in cord tissue, relatively lower number of CpGs (31%) showed hypomethylation in response to preterm in cord blood (Fig. 2b).
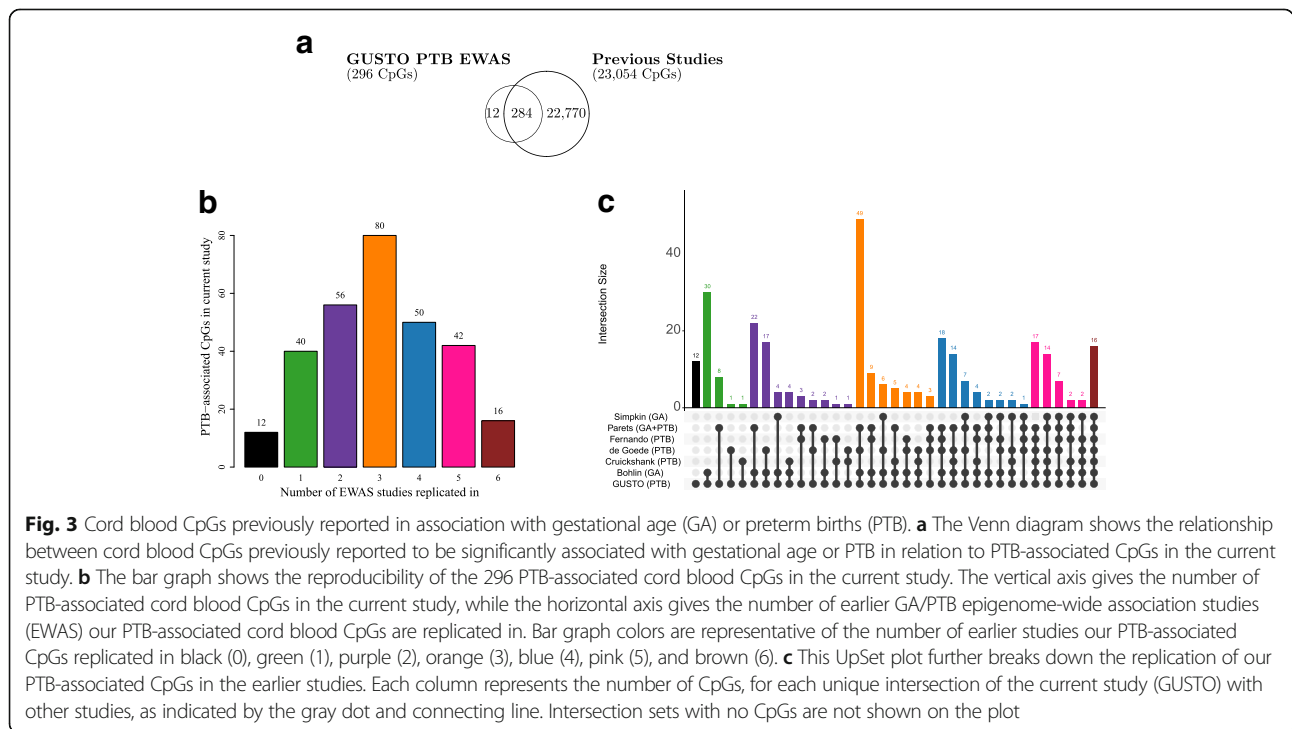
Majority of PTB-associated CpGs in cord blood are replicated in other PTB/GA EWAS

Since our study was conducted in an Asian population, we compared our cord blood EWAS findings with six previously reported studies in Caucasian and African American populations using the same Infinium HumanMethylation450 platform [7–12]. We consider a CpG to be replicated if it was reported in at least one of the previously conducted studies using the same Infinium HumanMethylation450 platform [7–12]. Of the 296 CpGs identified in our study, > 95% (284 CpGs) could be replicated in at least 1 of the previous studies and > 80% (244 CpGs) in at least 2 of the previous studies (Fig. 3, Additional file 2: Table S5 and S6), indicating robustness of the findings and a commonality in PTB associations across various ethnicities. Sixteen of these CpGs (12 genes) were reproducible in all 6 earlier independent studies, identifying robust epigenetic signatures of PTB. A subset of CpGs (22,770) from the previous 6 studies (Fig. 3) [7–12] were not identified in our study; however, these CpGs in general showed low reproducibility as only 15% of these were replicated in at least 1 of the studies and the remainder 85% showed no replication (Additional file 1: Figure S4). Nevertheless, the replicated CpGs from the earlier studies that are not

Fig. 2 Preterm births (PTB) were associated with global alterations in infants' cord blood DNA methylation. **a** Manhattan plot and **b** volcano plot illustrating the relationship of the 85,624 infant cord blood CpGs analyzed with respect to PTB. The top 10 CpGs with the smallest *P* values are indicated on both plots and labeled with the gene it is associated with or CpG identifier if the CpG lies within an intergenic region. Points on each plot represent individual CpGs which in **a** have genomic locations on the horizontal axis with alternating colors representing different chromosomes and in **b** have the change in DNA methylation *Z*-score on the horizontal axis. The red horizontal line in **a** represents the Bonferroni threshold ($P < 5.8 \times 10^{-7}$). Two hundred ninety-six infant cord blood CpGs were found to significantly associate with PTB and are indicated as red points in **b**. In both plots, the vertical axis represents the negative log10 *P* values with respect to preterm birth status, adjusted for infant sex, ethnicity, cell-type proportions, and bisulfite conversion batch

identified in our study are worthy of further consideration. We also applied the DNAm GA clocks published by Knight et al. [24] and Bohlin et al. [11] to predict GA in our study samples (Additional file 1: Figure S5). For both epigenetic clocks, the performance in cord blood (correlation = 0.52 for Knight et al. clock and correlation = 0.72 for Bohlin et al. clock, $n = 301$ term samples only) was better than the performance in cord tissue (correlation = 0.13 for Knight et al. clock and correlation = 0.16 for Bohlin et al. clock, $n = 951$ term samples only). The better performance in cord blood vs. cord tissue was not unexpected as both epigenetic clocks

Fig. 3 Cord blood CpGs previously reported in association with gestational age (GA) or preterm births (PTB). **a** The Venn diagram shows the relationship between cord blood CpGs previously reported to be significantly associated with gestational age or PTB in relation to PTB-associated CpGs in the current study. **b** The bar graph shows the reproducibility of the 296 PTB-associated cord blood CpGs in the current study. The vertical axis gives the number of PTB-associated cord blood CpGs in the current study, while the horizontal axis gives the number of earlier GA/PTB epigenome-wide association studies (EWAS) our PTB-associated cord blood CpGs are replicated in. Bar graph colors are representative of the number of earlier studies our PTB-associated CpGs replicated in black (0), green (1), purple (2), orange (3), blue (4), pink (5), and brown (6). **c** This UpSet plot further breaks down the replication of our PTB-associated CpGs in the earlier studies. Each column represents the number of CpGs, for each unique intersection of the current study (GUSTO) with other studies, as indicated by the gray dot and connecting line. Intersection sets with no CpGs are not shown on the plot

Wu *et al. Clinical Epigenetics*   (2019) 11:26

Page 5 of 12

were derived using (cord) blood samples. Similar to an earlier report [25], the performance improved when preterm infants were included, with better performance for cord blood (correlation = 0.69 for Knight et al. clock and correlation = 0.85 for Bohlin et al. clock, $n = 332$) than cord tissue (correlation = 0.15 for Knight et al. clock and correlation = 0.22 for Bohlin et al. clock, $n = 1019$).

DNA methylomes of cord blood and cord tissue respond differently to PTB

For CpGs significantly associated with PTB in at least one tissue, we also assessed whether there was evidence of tissue-dependent effects. For the 994 PTB-associated CpGs from cord tissue, 546 CpGs were removed from the cord blood dataset due to quality control filtering (426 of them were due to low inter-individual variation); for the remainder 448 CpGs, majority of the CpGs (143 at $P < 1e{-}4$, 310 at $P < 0.05$) showed evidence of tissue-dependent effects (Additional file 2: Table S7). Similarly for the 296 PTB-associated CpGs from cord blood, 102 CpGs were removed from the cord blood dataset due to quality control filtering (29 of them were due to low inter-individual variation); for the remainder 194 CpGs, majority of the CpGs (126 at $P < 1e{-}4$, 184 at $P < 0.05$) showed evidence of tissue-dependent effects (Additional file 2: Table S8).

DNAm status of genes affected by PTB in the two neonatal tissues represents distinct biological processes

We performed gene ontology analyses on the 994 cord tissue CpGs and 296 cord blood CpGs found to be significantly associated with PTB. Top gene ontology terms enriched with respect to cord tissue reflected biological processes primarily involved in fetal growth and development, i.e., Wnt signaling, bone remodeling, and extracellular matrix organization (Fig. 4, Additional file 1: Figure S6 and S7, Additional file 2: Table S9). In contrast, cord blood reflected regulation of T cell differentiation, inositol lipid-mediated signaling, and regulation of RNA stability (Fig. 4, Additional file 1: Figure S8 and S9, Additional file 2: Table S10). These results are consistent with the fact that variable CpGs in each tissue tend to over-represent certain pathways. This also is clearly evident from the gene ontology analysis of all variable CpGs identified on the Infinium HumanMethylation450 platform from each tissue (Additional file 1: Figure S10 and S11).

Majority of PTB-associated CpGs in cord tissue and cord blood were also associated with GA

Lastly, we also examined the associations between GA and DNAm in each tissue. In this analysis, GA was modeled as a continuous variable instead of a binary variable (preterm vs. term). After adjustment for multiple testing, 4075 CpGs ($P < 3.7e{-}7$) were significantly associated with GA in cord tissue (Additional file 2: Table S11). Upon analysis using cord blood, 1916 CpGs ($P < 5.8e{-}7$) were associated with GA (Additional file 2: Table S12), 94 of these overlapped with the 4075 cord tissue GA-associated CpGs. Comparison of GA-associated vs. PTB-associated CpGs (Additional file 1: Figure S12) showed that > 95% of the 994 PTB-associated CpGs in cord tissue were also GA-associated (950 with $P < 3.7e{-}7$ and 993 with $P < 1e{-}4$ in an analysis using GA). Similarly, most of the 296 PTB-associated CpGs in cord blood remained GA-associated (284 with $P < 5.8e{-}7$ and 293 with $P < 1e{-}4$). These results suggests PTB-associated CpGs may also be a signature of GA. Gene ontology
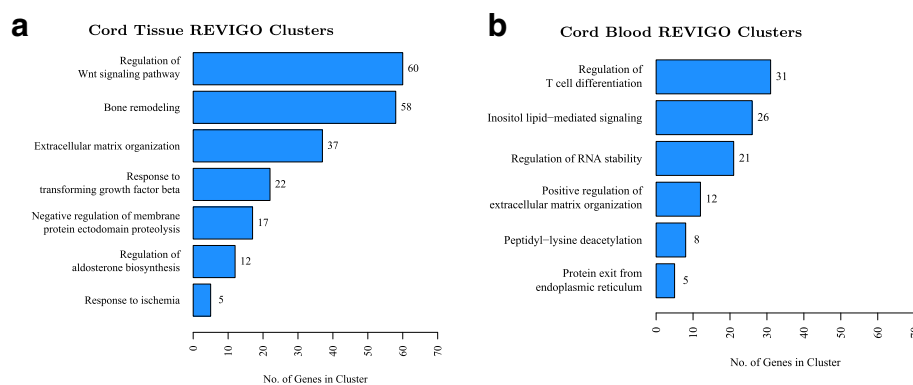

Fig. 4 a, **b** REVIGO summarized Gene Ontology Clusters with respect to preterm birth (PTB)-associated CpGs in both cord tissue and cord blood. Gene ontology (GO) enrichment was performed on PTB-associated CpGs in both cord tissue and cord blood for each tissue separately using missMethyl. REVIGO was then used to reclassify the biological process-related enriched GO terms (parent GO term containing under 300 genes, semantic similarity measure between each GO term < 0.7). Cord tissue CpGs had 10 GO clusters from 41 unique GO terms, while cord blood CpGs had 10 GO clusters from 43 unique GO terms. GO clusters with 5 or more genes are represented by the bar graphs, with plots on the left and right corresponding to cord tissue and cord blood respectively. The vertical axis of the bar graphs represents the REVIGO cluster names, while the horizontal axis represents the number of genes in the REVIGO cluster containing at least one significantly associated CpG

Wu *et al. Clinical Epigenetics*   (2019) 11:26

Page 6 of 12

analyses performed on the 4075 cord tissue GA-associated CpGs and 1916 cord blood GA-associated CpGs gave similar conclusions as the analyses performed on PTB-associated CpGs. Specifically, cord tissue CpGs showed enrichment of pathways (Additional file 1: Figure S13) related to fetal growth and development (Additional file 1: Figure S14, Additional file 2: Table S13), while cord blood CpGs showed enrichment of immune response pathways (Additional file 1: Figure S15, Additional file 2: Table S14). We also compared these 1916 cord blood GA-associated CpGs with those reported by previous studies [7–12]. Of the 1916 cord blood GA-associated CpGs identified in the current study, 89% (1714 CpGs) could be replicated in at least 1 of the previous studies and 60% (1141 CpGs) in at least 2 of the previous studies (Additional file 1: Figure S16, Additional file 2: Table S15). However, the replication of the 296 cord blood PTB-associated CpGs with previous studies was relatively higher as > 95% of these CpGs replicated in at least 1 of the previous studies and > 80% replicated in at least 2 other studies.

Discussion

In this study, we report associations with DNAm profiles in neonates born preterm by using tissues of different germinal origins, i.e., cord tissue and cord blood. The key findings from our study include (1) the replication of PTB/GA-associated cord blood CpGs across different studies and ethnicities to identify robust epigenetic signatures of PTB, (2) the identification of DNAm associations with PTB in cord tissue, and (3) the importance of evaluating the DNA methylomes of two germinally distinct neonatal tissues to capture a more comprehensive view of the molecular pathways associated with PTB.

Replication of CpGs associated with GA/PTB in cord blood across different studies and ethnicities

More than 95% of the CpGs identified in our cord blood PTB EWAS were replicated in previous PTB/GA EWAS studies [7–12]. In particular, cg23062810 from *CLIP2* gene was replicated across six independent studies. *CLIP2* gene also seems to be a hotspot for PTB/GA-associated DNAm changes, as 6 additional CpGs have been previously reported from this gene—cg16356456 [7–12], cg04952324 [8–11], cg11573518 [11], cg02935052 [11], cg21375204 [11], and cg19501108 [10]. Notably, 2 CpGs adjoining cg23062810, i.e., cg16356456 and cg11573518, also showed moderate significance (P value $< 10^{-5}$) in our study. The CpG trio of cg23062810, cg16356456, and cg11573518 is a promising candidate epigenetic signature for functional studies, as they are not only consistently reported to be hypermethylated in cord blood of preterm neonates, but also span a short 224-bp genomic region containing DNaseI hypersensitive site and several known transcription factor binding sites. CLIP2 is a cytoplasmic

linker protein expressed in the brain [26], with its haploinsufficiency linked to motor coordination abnormalities [27]. *CLIP2* deletion is linked to Williams-Beuren syndrome, but deletion of a single copy alone is insufficient to result in the physical or cognitive characteristics of the disease [28].

Furthermore, in spite of the interrogation of PTB associations in an Asian population within our study, we achieved robust replication of 16 CpGs across all 6 earlier PTB/GA EWAS studies conducted in other populations of Caucasian/African American origin. These 16 CpGs span 12 genes, with 4 of these genes containing at least 2 PTB-associated CpGs in the current study. These genes include interleukin 21 receptor (*IL21R*), a key component of the adaptive immune system [29]; *NCOR2*, a relatively ubiquitously expressed repressor linked to a wide variety of biological processes including metabolism, inflammation, and circadian rhythm [30]; proline-rich 5 like (*PRR5L*), involved in the cellular response to oxidative stress [31]; and insulin-like growth factor 2 mRNA-binding protein 1 (*IGF2BP1*), a tightly regulated cell proliferation protein highly expressed during embryogenesis [32, 33]. Notably, the *PRR5L* gene carries 10 previously reported GA/PTB-associated CpGs, 3 of which we found to be PTB-associated in the current study (cg08943494, cg00220721, cg22117805). Although the exact function of PRR5L with respect to pregnancy is unknown, PRR5L suppresses a key regulator of cellular mTORC2 in vitro, which in turn is regulated by lysophosphatidic acid (LPA) and Gα12 activity [34]. LPA is implicated in the maintenance of pregnancy [35], uterine contractility [36], and infection-related preterm labor [37]; while Gα12 is a molecular regulator of extracellular stimuli, including oxidative stress [38]. There is also emerging evidence that mTOR-related genes are differentially expressed between term and preterm labor as well as between labor and non-labor myometrial [39].

Identification of associations between DNAm and PTB in cord tissue

In addition to the findings from cord blood, we identified 994 CpGs to significantly associate with PTB in cord tissue, of which only 10 CpGs overlapped with cord blood CpGs. Our cord tissue findings provide new insights into the epigenetic landscape of neonates born preterm as this birth tissue has not been explored in this context before. Most importantly, the analysis of two neonatal tissues representing different cell type lineages provides a wider coverage of biological processes associated with PTB.

Combination of EWAS in two neonatal tissues captures a comprehensive view of the molecular pathways associated with PTB

The two neonatal tissues gave deeper insights into the plausible molecular pathways associated with PTB.

Gene networks in cord blood indicated the role of inflammation in PTB, which is in agreement with the previous findings implicating the role of inflammation in the etiology of PTB [6, 40]. The top most statistically significant PTB-CpGs from cord blood were found in genes involved in inflammation such as *TRAF5*, a key regulator of both canonical (via TNFα [41]) and non-canonical (via lymphotoxins [42]) NF-kappaB activation, and *MYLK*, a relatively ubiquitously expressed gene implicated in several inflammatory diseases [43] and also the main target for oxytocin-induced phosphorylation, downregulation of which follows uterine contraction at term [44]. Immune-related genes that were highly reproduced across different studies include *NCOR2*, an integral corepressor within the Notch signaling [45] with links to NK-kappaB-mediated apoptosis [46]; zinc finger and BTB domain containing 7B (*ZBTB7B*), a key regulator of CD4+T cell commitment [47]; PDZ and LIM domain protein 2 (*PDLIM2*), a key inhibitor of inflammatory response through NF-kappaB [48]; and *IL2RA*, a key component in immunological function primarily through the establishment of T cell immunological memory [49].

Gene ontology terms linked with cord blood CpGs also reflected the dominance of immune-related biological processes despite the adjustment for cellular heterogeneity. The largest gene ontology cluster enriched in the pathway analysis was regulation of T cell differentiation, a hallmark of innate immune system development, which includes genes such as tripartite motif containing 22 (*TRIM22*, interferon signaling [50]), interleukin 1 receptor-associated kinase 2 (*IRAK2*, inflammatory response to infection [51]), and caspase recruitment domain family member 11 (*CARD11*, critical component of T cell and B cell signaling [52]). The next two largest clusters also featured several gene ontology terms with various immune-related nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kappaB) components.

The role of immune-related genes is also apparent, albeit to a smaller degree, in cord tissue CpGs significantly associated with PTB. Prominent examples include *NKFBIA*, which binds to the nuclear localization signal of the inflammatory response element NF-kappa-B/REL complex, preventing transcription and inflammatory response [53]; and *NFIL3* (nuclear factor, interleukin 3 activated), a transcription regulator, mostly inhibiting many genes [54], but also known to activate interleukin-3 [55], mediating pro-B lymphocyte survival [56]. Incidentally, NFKBIA appears to be upregulated in placenta with history of chorioamnionitis, as well as those complicated by preterm premature rupture of membrane (PPROM) cases [57]. Upregulation of NFKBIA is suggested to be a form of anti-inflammatory response to inflammatory insults [57].

NFIL3, on the other hand, is downregulated with term within CD34+ cord blood fractions [58], consistent with the comparatively larger population of immature hematopoietic progenitor population in preterm cord blood. Collectively, differentially methylated CpGs from both tissues highlight the role of inflammatory genes in PTB, with a larger representation in cord blood than cord tissue.

Cord tissue CpGs significantly associated with PTB were found in genes with more diverse gene functions as opposed to primarily immune responses seen in cord blood. These included general transcription factor genes such as protein C-Ets-2 (*ETS2*) [59] and specificity protein 1 (*SP1*) [60]. Incidentally, ETS2 was previously reported to be downregulated in preterm placentas with spontaneous labor [61], while differentially expressed genes with respect to peripheral blood in mothers who delivered preterm possessed over-representation of SP1 binding sites within their promoters [62]. Gene ontology enrichment analysis revealed cord tissue CpGs to mostly lie in genes related to physiological growth and development. In particular, bone development was the largest grouped cord tissue gene ontology result, including genes such as parathyroid hormone 1 receptor (*PTH1R*, surface receptor of osteoblasts [63]), bone morphogenetic proteins 2 and 6 (*BMP2*, *BMP6*, simulator of bone growth [64]), and matrix metallopeptidase 7 (*MMP7*, associated with bone remodeling [65]). This was followed by regulation of Wnt signaling pathway—a pathway which plays a central role in embryonic development [66], with members such as Wnt family members 8A and 11 (*WNT8A*, involved in axis patterning [67]; *WNT11* is involved in the skeletal, kidney, and lung development [68]). The third largest gene cluster was related to extracellular matrix (ECM) organization. ECM impacts a number of cellular functions critical for normal fetal development and morphogenesis [69]. The most familiar developmental function attributed to ECM is cell migration during fetal development and organogenesis that is facilitated by cycles of cell adhesion and deadhesion. ECM also plays important structural roles in defining tissue boundaries, branching morphogenesis, developing tissue asymmetry and growth factor signaling.

Study limitations

This study has a few limitations. First, while we have observed global DNAm alterations in the two neonatal tissues at CpG sites assayed using the Infinium Human-Methylation450 platform, the CpGs assayed by the platform were not randomly selected from the DNA methylome. Consequently, it is unclear how the findings will extend to the rest of the DNA methylome. Second, as supported by our findings here and in a

Wu *et al. Clinical Epigenetics*   (2019) 11:26

Page 8 of 12

previous publication [23], EWAS conducted using different tissues can give very distinct findings. While use of clinically available tissues like cord tissue and cord blood is convenient, use of these two tissues may not completely mirror the effects of PTB in target tissues. Thus, further research is necessary to investigate if these findings can be extrapolated to the relevant tissues of interest. Third, while we have successfully replicated our findings in cord blood using results from published literature, due to the lack of availability of cord tissue DNAm data, we are unable to replicate the cord tissue findings in an independent cohort. However, the robust (> 95% CpGs) replication of the findings in previously reported PTB studies in cord blood and the use of larger sample size suggest that our novel cord tissue findings are likely to be robust too.

Conclusion

Using DNAm profiles from two different neonatal tissues (cord tissue and cord blood), we provide the epigenetic status of a broader spectrum of molecular pathways associated with PTB. Our findings suggest that genes involved in inflammation and fetal developmental processes play a key role in PTB. Further research is necessary to identify the specific role played by these epigenetic changes on the postnatal developmental and health trajectories of the offspring.

Methods

Study population

Between June 2009 and September 2010, healthy pregnant women were recruited in their first trimester of pregnancy from two major public hospitals in Singapore, namely the KK Women's and Children's Hospital (KKH) and the National University Hospital (NUH), to participate in the Growing Up in Singapore Towards Healthy Outcomes (GUSTO) birth cohort study [70]. To participate in the study, pregnant women had to satisfy the following inclusion criteria: (1) be of at least 18 years of age; (2) hold Singapore citizenship or permanent residency, or intent to reside in Singapore for the next 5 years; (3) be of Chinese, Malay, or Indian ethnic origin, confirmed through homogeneous parental ethnic background and genotyping; (4) intent to deliver at either NUH or KKH; and (5) intent to donate cord tissue and cord blood. The exclusion criteria included (1) women on chemotherapy, (2) women with significant health conditions such as type 1 diabetes mellitus and psychosis, and (3) women on specific medications such as psychotropic drugs. The present analysis was restricted to live singleton births with infant DNAm data (cord tissue or cord blood).

Determining GA, infant sex, and ethnicity

GA was determined by ultrasonography in the first trimester of pregnancy. PTB was defined as GA < 37 weeks. Child sex was extracted from the medical records. Ethnicity was self-reported by the mother at study recruitment.

Tissue collection and processing

Detailed information on cord tissue and cord blood collection as well as processing has been previously described [23]. Briefly, cord blood was collected post-delivery by either dripping the blood in EDTA tubes for normal deliveries or collecting via a syringe in the event of assisted deliveries. Collected cord blood was centrifuged at 4 °C, 3000g for 5 min, and the buffy coat extracted was stored at − 80 °C until subsequent DNA extraction. DNA extraction of cord blood was carried out using QIAsymphony DNA Kit as per the manufacturer's instructions. After collection of the cord blood, cord tissue was cleaned with phosphate buffer saline (PBS) solution. The cord was then snap-frozen in liquid nitrogen and stored at − 80 °C until subsequent DNA extraction. Before DNA extraction, frozen umbilical cords were crushed using a mortar and pestle, treated with 10 U/mL hydraluronidase enzyme and homogenized using a Xiril Dispomix Homogeniser. Proteinase K was added to the homogenate and incubated overnight at 55 °C. Cord tissue DNA was then extracted as described earlier [23].

DNAm profiling and data processing

DNA methylomes for cord tissue and cord blood were profiled and processed separately using the Infinium HumanMethylation450 platform (Additional files 3 and 4). Data processing was conducted using an in-house quality control procedure that was previously described [71]. Briefly, we exported raw DNAm beta values from GenomeStudio™ and set probes with less than three beads for either the methylated or unmethylated channel or with detection P value > 0.01 to missing. We then performed color adjustment and normalization of the type 1 and 2 probes and excluded sex chromosome probes. As part of the study design for DNAm profiling, samples were randomized across chip and position on chip with respect to key variables including GA, infant sex, and ethnicity. Thus, expectedly, PTB did not associate with chip or position effects. For both tissues, a principal component analysis of the raw DNAm revealed chip to associate most significantly with the raw DNAm data. DNAm data for both tissues were thus adjusted for chip using COMBAT, removing CpGs with missing values across all 12 positions on any chip [72]. For the remainder technical variables that were associated with top principal components of the DNAm data, but were not randomized, PTB was associated with bisulfite conversion batch (both cord tissue and cord blood) and DNA extraction batch (cord tissue only), and

these variables were included as covariates in all regression models. Finally, cross-hybridizing probes [73, 74], CpGs on or within a single-base extension of a SNP and CpGs with multi-modal distributions were excluded from the analysis. As CpGs with low inter-individual variation in each tissue may be more reflective of the technical variation than true biological signal, to reduce false positives and increase overall study power [75, 76], we further excluded CpGs that had low inter-individual variation in each tissue (i.e., DNAm range under 10% or DNAm of the 99th centile minus 1st centile under 5%). After quality control and exclusion of CpGs with low variability, 134,676 CpGs (cord tissue) and 85,624 CpGs (cord blood) were available for subsequent analysis. For infant cord tissue, cellular proportions for stromal, endothelial, epithelial, and blood were estimated using a reference panel [77] and their principal components were adjusted as covariates in all regression models. Likewise, for infant cord blood, cell-type proportions for granulocytes, monocytes, natural killer cells, B cells, CD4+ T cells and CD8+ T cells were estimated using a reference panel [77] and their principal components were adjusted as covariates in all regression models. CpGs were annotated with respect to gene features (promoter, 5′-UTR, exon, intron, 3′-UTR, TTS, and intergenic regions) using Homer *annotatePeaks* function (hg19).

Statistical analysis
Association between DNAm and PTB
To examine the association between DNAm and PTB for each tissue, we fitted a linear regression model with DNAm as the dependent variable and PTB as the independent variable, adjusted for technical variables that associated with PTB (bisulfite conversion batch and DNA extraction batch), infant sex, ethnicity, and estimated cell-type proportions. Infant sex and ethnicity were selected as covariates for inclusion in the regression models based on a priori evidence of their playing key roles in DNAm and/or PTB. For each CpG, individuals with outlier DNAm values (defined as DNAm values exceeding the cohort median ± twice the interquartile range for each CpG) were excluded from the analysis. PTB was coded as a binary variable, with 1 = term and 0 = preterm; thus, a negative regression coefficient implies that DNAm levels were generally higher among the preterm infants compared to term infants.

For CpGs significantly associated with PTB in at least one tissue, we also assessed whether there was evidence of tissue-dependent effects. This analysis was performed by fitting a general linear model with an unstructured covariance structure to a combined dataset with DNAm data from both tissues, including main effect terms for tissue and PTB, and an interaction term between PTB and tissue and other covariates. The interaction term between PTB and tissue provides an estimate of the difference in PTB-DNAm association in the two tissues, and a statistical test of this interaction term provides a formal test of tissue-dependent effects.

Pathway analysis
For genes where the CpGs were significantly associated with PTB after adjustment for multiple testing using a Bonferroni correction, we further examined them for enrichment of gene ontology biological pathways using the *gometh* function in the MissMethyl R package [78], which maps CpG sites to their nearest gene and corrects for bias due to non-uniform coverage of genes on the Infinium HumanMethylation450 array. To consolidate and summarize the pathway enrichment analysis results from *gometh*, nominally significant GO terms ($P < 0.01$) within the "biological processes" category were further run through the REVIGO tool, which avoids reporting GO terms with greater than 70% in semantic similarity measure [79]. As GO terms involving many genes may not inform precise gene functionalities, larger GO terms (containing 300 or more genes) were removed before running REVIGO. The results from REVIGO were visualized using TreeMaps.

Sensitivity analysis
We also conducted sensitivity analyses where we further adjusted for mode of delivery, maternal hypertension, maternal age, smoking, parity, and position on chip (sensitivity analysis 1). To further allow for the possibility of unmeasured technical artifacts or un-accounted cell-type proportions, we also used surrogate variable analysis (SVA) to directly estimate sources of batch effects and/or cell-type composition from the DNAm data. The resulting estimated surrogate variables from the SVA could potentially capture both batch effects and cell-type composition. We conducted additional sensitivity analyses (sensitivity analysis 2), where we repeated the association analyses between PTB and DNAm, adjusting for surrogate variables from the SVA, on top of infant sex and ethnicity [80, 81].

Comparison of PTB-associated CpGs in cord blood with previously published studies
We compared our cord blood PTB EWAS findings with PTB/GA EWAS findings from previous studies [7–12]. For a fair comparison, we restricted this analysis to the studies conducted using the same Infinium Human-Methylation450 platform. We also applied the DNAm GA clocks published by Knight et al. [24] and Bohlin et al. [11] to predict GA in our study samples. The clocks published by Knight et al. and Bohlin et al. were applied to our cord tissue and cord blood DNAm data separately. For this analysis, raw DNAm data without any processing or quality control filtering was used.

Wu *et al. Clinical Epigenetics* (2019) 11:26

Page 10 of 12

Associations between DNAm and GA

Since a number of previous EWAS were conducted using GA as a continuous variable instead of PTB as a binary variable, we also conducted an additional analysis using GA as a continuous variable. For each tissue, we fitted a linear regression model with DNAm as the dependent variable and GA as the independent variable, adjusted for the same covariates as before. Pathway analysis and comparison with earlier reports were performed similarly.

Additional files

Additional file 1: Supplementary figures. (PDF 6638 kb)

Additional file 2: Supplementary tables. (ZIP 3966 kb)

Additional file 3: DNA methylation data for cord blood. (PHENO 2410 kb)

Additional file 4: DNA methylation data for cord tissue. (PHENO 1130000 kb)

Author details
[1]Singapore Institute for Clinical Sciences, A*STAR, 30 Medical Drive, Singapore 117609, Singapore. [2]Duke-NUS Medical School, Singapore, Singapore. [3]Department of Medical Genetics, Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, University of British Columbia, Vancouver, Canada. [4]KK Women's and Children's Hospital, Singapore, Singapore. [5]Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore. [6]Centre for Human Evolution, Adaptation and Disease, Liggins Institute, University of Auckland, Auckland, New Zealand. [7]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore.

References
1. March of Dimes, PMNCH, Save the Children, WHO. Born too soon: the global action report on preterm birth. Geneva: World Health Organization; 2012.
2. Preterm birth. https://www.who.int/pmnch/media/news/2012/201204_borntoosoon-report.pdf.
3. Huddy CL, Johnson A, Hope PL. Educational and behavioural problems in babies of 32-35 weeks gestation. Arch Dis Child Fetal Neonatal Ed. 2001; 85(1):F23–8.
4. Wang ML, Dorer DJ, Fleming MP, Catlin EA. Clinical outcomes of near-term infants. Pediatrics. 2004;114(2):372–6.
5. Beck S, Wojdyla D, Say L, Betran AP, Merialdi M, Requejo JH, Rubens C, Menon R, Van Look PF. The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. Bull World Health Organ. 2010;88(1):31–8.
6. Muglia LJ, Katz M. The enigma of spontaneous preterm birth. N Engl J Med. 2010;362(6):529–35.
7. Fernando F, Keijser R, Henneman P, van der Kevie-Kersemaekers AM, Mannens MM, van der Post JA, Afink GB, Ris-Stalpers C. The idiopathic preterm delivery methylation profile in umbilical cord blood DNA. BMC Genomics. 2015;16:736.
8. Cruickshank MN, Oshlack A, Theda C, Davis PG, Martino D, Sheehan P, Dai Y, Saffery R, Doyle LW, Craig JM. Analysis of epigenetic changes in survivors of preterm birth reveals the effect of gestational age and evidence for a long term legacy. Genome Med. 2013;5(10):96.
9. Parets SE, Conneely KN, Kilaru V, Fortunato SJ, Syed TA, Saade G, Smith AK, Menon R. Fetal DNA methylation associates with early spontaneous preterm birth and gestational age. PLoS One. 2013;8(6):e67489.

Wu *et al. Clinical Epigenetics*   (2019) 11:26

Page 11 of 12

10. de Goede OM, Lavoie PM, Robinson WP. Cord blood hematopoietic cells from preterm infants display altered DNA methylation patterns. Clin Epigenetics. 2017;9:39.

11. Bohlin J, Haberg SE, Magnus P, Reese SE, Gjessing HK, Magnus MC, Parr CL, Page CM, London SJ, Nystad W. Prediction of gestational age based on genome-wide differentially methylated regions. Genome Biol. 2016; 17(1):207.

12. Simpkin AJ, Suderman M, Gaunt TR, Lyttleton O, McArdle WL, Ring SM, Tilling K, Davey Smith G, Relton CL. Longitudinal analysis of DNA methylation associated with birth weight and gestational age. Hum Mol Genet. 2015;24(13):3752–63.

13. Menon R, Conneely KN, Smith AK. DNA methylation: an epigenetic risk factor in preterm birth. Reprod Sci. 2012;19(1):6–13.

14. 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(9):784–92.

15. Burris HH, Braun JM, Byun HM, Tarantini L, Mercado A, Wright RJ, Schnaas L, Baccarelli AA, Wright RO, Tellez-Rojo MM. Association between birth weight and DNA methylation of IGF2, glucocorticoid receptor and repetitive elements LINE-1 and Alu. Epigenomics. 2013;5(3):271–81.

16. Burris HH, Rifas-Shiman SL, Baccarelli A, Tarantini L, Boeke CE, Kleinman K, Litonjua AA, Rich-Edwards JW, Gillman MW. Associations of LINE-1 DNA methylation with preterm birth in a prospective cohort study. J Dev Orig Health Dis. 2012;3(3):173–81.

17. Mitsuya K, Singh N, Sooranna SR, Johnson MR, Myatt L. Epigenetics of human myometrium: DNA methylation of genes encoding contraction-associated proteins in term and preterm labor. Biol Reprod. 2014;90(5):98.

18. Sparrow S, Manning JR, Cartier J, Anblagan D, Bastin ME, Piyasena C, Pataky R, Moore EJ, Semple SI, Wilkinson AG, et al. Epigenomic profiling of preterm infants reveals DNA methylation differences at sites associated with neural function. Transl Psychiatry. 2016;6:e716.

19. Liu Y, Hoyo C, Murphy S, Huang Z, Overcash F, Thompson J, Brown H, Murtha AP. DNA methylation at imprint regulatory regions in preterm birth and infection. Am J Obstet Gynecol. 2013;208(5):395 e391–397.

20. Behnia F, Parets SE, Kechichian T, Yin H, Dutta EH, Saade GR, Smith AK, Menon R. Fetal DNA methylation of autism spectrum disorders candidate genes: association with spontaneous preterm birth. Am J Obstet Gynecol. 2015;212(4):533 e531–539.

21. Schroeder JW, Conneely KN, Cubells JC, Kilaru V, Newport DJ, Knight BT, Stowe ZN, Brennan PA, Krushkal J, Tylavsky FA, et al. Neonatal DNA methylation patterns associate with gestational age. Epigenetics. 2011; 6(12):1498–504.

22. Lee H, Jaffe AE, Feinberg JI, Tryggvadottir R, Brown S, Montano C, Aryee MJ, Irizarry RA, Herbstman J, Witter FR, et al. DNA methylation shows genome-wide association of NFIX, RAPGEF2 and MSRB3 with gestational age at birth. Int J Epidemiol. 2012;41(1):188–99.

23. Lin X, Teh AL, Chen L, Lim IY, Tan PF, MacIsaac JL, Morin AM, Yap F, Tan KH, Saw SM, et al. Choice of surrogate tissue influences neonatal EWAS findings. BMC Med. 2017;15(1):211.

24. Knight AK, Craig JM, Theda C, Bækvad-Hansen M, Bybjerg-Grauholm J, Hansen CS, Hollegaard MV, Hougaard DM, Mortensen PB, Weinsheimer SM, et al. An epigenetic clock for gestational age at birth based on blood methylation data. Genome Biol. 2016;17(1):206.

25. Simpkin AJ, Suderman M, Howe LD. Epigenetic clocks for gestational age: statistical and study design considerations. Clin Epigenetics. 2017;9:100.

26. Hoogenraad CC, Eussen BH, Langeveld A, van Haperen R, Winterberg S, Wouters CH, Grosveld F, De Zeeuw CI, Galjart N. The murine CYLN2 gene: genomic organization, chromosome localization, and comparison to the human gene that is located within the 7q11.23 Williams syndrome critical region. Genomics. 1998;53(3):348–58.

27. van Hagen JM, van der Geest JN, van der Giessen RS, Lagers-van Haselen GC, Eussen HJ, Gille JJ, Govaerts LC, Wouters CH, de Coo IF, Hoogenraad CC, et al. Contribution of CYLN2 and GTF2IRD1 to neurological and cognitive symptoms in Williams syndrome. Neurobiol Dis. 2007;26(1):112–24.

28. Vandeweyer G, Van der Aa N, Reyniers E, Kooy RF. The contribution of CLIP2 haploinsufficiency to the clinical manifestations of the Williams-Beuren syndrome. Am J Hum Genet. 2012;90(6):1071–8.

29. Johnson LD, Jameson SC. Immunology. A chronic need for IL-21. Science. 2009;324(5934):1525–6.

30. Mottis A, Mouchiroud L, Auwerx J. Emerging roles of the corepressors NCoR1 and SMRT in homeostasis. Genes Dev. 2013;27(8):819–35.

31. Holmes B, Artinian N, Anderson L, Martin J, Masri J, Cloninger C, Bernath A, Bashir T, Benavides-Serrato A, Gera J. Protor-2 interacts with tristetraprolin to regulate mRNA stability during stress. Cell Signal. 2012;24(1):309–15.

32. Bell JL, Wachter K, Muhleck B, Pazaitis N, Kohn M, Lederer M, Huttelmaier S. Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? Cell Mol Life Sci. 2013;70(15):2657–75.

33. Nielsen J, Christiansen J, Lykke-Andersen J, Johnsen AH, Wewer UM, Nielsen FC. A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development. Mol Cell Biol. 1999;19(2):1262–70.

34. Gan X, Wang J, Wang C, Sommer E, Kozasa T, Srinivasula S, Alessi D, Offermanns S, Simon MI, Wu D. PRR5L degradation promotes mTORC2-mediated PKC-delta phosphorylation and cell migration downstream of Galpha12. Nat Cell Biol. 2012;14(7):686–96.

35. Tokumura A, Kanaya Y, Miyake M, Yamano S, Irahara M, Fukuzawa K. Increased production of bioactive lysophosphatidic acid by serum lysophospholipase D in human pregnancy. Biol Reprod. 2002;67(5):1386–92.

36. Tokumura A, Fukuzawa K, Yamada S, Tsukatani H. Stimulatory effect of lysophosphatidic acids on uterine smooth muscles of non-pregant rats. Arch Int Pharmacodyn Ther. 1980;245(1):74–83.

37. Ye X, Chun J. Lysophosphatidic acid (LPA) signaling in vertebrate reproduction. Trends Endocrinol Metab. 2010;21(1):17–24.

38. Cho MK, Kim WD, Ki SH, Hwang JI, Choi S, Lee CH, Kim SG. Role of Galpha12 and Galpha13 as novel switches for the activity of Nrf2, a key antioxidative transcription factor. Mol Cell Biol. 2007;27(17):6195–208.

39. Foster HA, Davies J, Pink RC, Turkcigdem S, Goumenou A, Carter DR, Saunders NJ, Thomas P, Karteris E. The human myometrium differentially expresses mTOR signalling components before and during pregnancy: evidence for regulation by progesterone. J Steroid Biochem Mol Biol. 2014;139:166–72.

40. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. Lancet. 2008;371(9606):75–84.

41. Kawamata S, Hori T, Imura A, Takaori-Kondo A, Uchiyama T. Activation of OX40 signal transduction pathways leads to tumor necrosis factor receptor-associated factor (TRAF) 2- and TRAF5-mediated NF-kappaB activation. J Biol Chem. 1998;273(10):5808–14.

42. Nakano H, Oshima H, Chung W, Williams-Abbott L, Ware CF, Yagita H, Okumura K. TRAF5, an activator of NF-kappaB and putative signal transducer for the lymphotoxin-beta receptor. J Biol Chem. 1996;271(25):14661–4.

43. Xiong Y, Wang C, Shi L, Wang L, Zhou Z, Chen D, Wang J, Guo H. Myosin light chain kinase: a potential target for treatment of inflammatory diseases. Front Pharmacol. 2017;8:292.

44. Salomonis N, Cotte N, Zambon AC, Pollard KS, Vranizan K, Doniger SW, Dolganov G, Conklin BR. Identifying genetic networks underlying myometrial transition to labor. Genome Biol. 2005;6(2):R12.

45. Kao HY, Ordentlich P, Koyano-Nakagawa N, Tang Z, Downes M, Kintner CR, Evans RM, Kadesch T. A histone deacetylase corepressor complex regulates the Notch signal transduction pathway. Genes Dev. 1998; 12(15):2269–77.

46. Hoberg JE, Yeung F, Mayo MW. SMRT derepression by the IkappaB kinase alpha: a prerequisite to NF-kappaB transcription and survival. Mol Cell. 2004; 16(2):245–55.

47. Wang L, Wildt KF, Castro E, Xiong Y, Feigenbaum L, Tessarollo L, Bosselut R. The zinc finger transcription factor Zbtb7b represses CD8-lineage gene expression in peripheral CD4+ T cells. Immunity. 2008;29(6):876–87.

48. Tanaka T, Grusby MJ, Kaisho T. PDLIM2-mediated termination of transcription factor NF-kappaB activation by intranuclear sequestration and degradation of the p65 subunit. Nat Immunol. 2007;8(6):584–91.

49. Liao W, Lin JX, Leonard WJ. IL-2 family cytokines: new insights into the complex roles of IL-2 as a broad regulator of T helper cell differentiation. Curr Opin Immunol. 2011;23(5):598–604.

50. Tissot C, Mechti N. Molecular cloning of a new interferon-induced factor that represses human immunodeficiency virus type 1 long terminal repeat expression. J Biol Chem. 1995;270(25):14891–8.

51. Wan Y, Xiao H, Affolter J, Kim TW, Bulek K, Chaudhuri S, Carlson D, Hamilton T, Mazumder B, Stark GR, et al. Interleukin-1 receptor-associated kinase 2 is critical for lipopolysaccharide-mediated post-transcriptional control. J Biol Chem. 2009;284(16):10367–75.

52. Chan W, Schaffer TB, Pomerantz JL. A quantitative signaling screen identifies CARD11 mutations in the CARD and LATCH domains that induce Bcl10 ubiquitination and human lymphoma cell survival. Mol Cell Biol. 2013;33(2):429–43.

53. Scherer DC, Brockman JA, Chen Z, Maniatis T, Ballard DW. Signal-induced degradation of I kappa B alpha requires site-specific ubiquitination. Proc Natl Acad Sci U S A. 1995;92(24):11259–63.

54. Keniry M, Dearth RK, Persans M, Parsons R. New frontiers for the NFIL3 bZIP transcription factor in cancer, metabolism and beyond. Discoveries (Craiova). 2014;2(2):e15.

55. Zhang W, Zhang J, Kornuc M, Kwan K, Frank R, Nimer SD. Molecular cloning and characterization of NF-IL3A, a transcriptional activator of the human interleukin-3 promoter. Mol Cell Biol. 1995;15(11):6055–63.

56. Ikushima S, Inukai T, Inaba T, Nimer SD, Cleveland JL, Look AT. Pivotal role for the NFIL3/E4BP4 transcription factor in interleukin 3-mediated survival of pro-B lymphocytes. Proc Natl Acad Sci U S A. 1997;94(6):2609–14.

57. Kumar N, Nandula P, Menden H, Jarzembowski J, Sampath V, Placental TLR. NLR expression signatures are altered with gestational age and inflammation. J Matern Fetal Neonatal Med. 2017;30(13):1588–95.

58. Podesta M, Bruschettini M, Cossu C, Sabatini F, Dagnino M, Romantsik O, Spaggiari GM, Ramenghi LA, Frassoni F. Preterm cord blood contains a higher proportion of immature hematopoietic progenitors compared to term samples. PLoS One. 2015;10(9):e0138680.

59. Sumarsono SH, Wilson TJ, Tymms MJ, Venter DJ, Corrick CM, Kola R, Lahoud MH, Papas TS, Seth A, Kola I. Down's syndrome-like skeletal abnormalities in Ets2 transgenic mice. Nature. 1996;379(6565):534–7.

60. Zhao C, Meng A. Sp1-like transcription factors are regulators of embryonic development in vertebrates. Develop Growth Differ. 2005;47(4):201–11.

61. Eidem HR, Rinker DC, WET A, Buhimschi IA, Buhimschi CS, Dunn-Fletcher C, Kallapur SG, Pavlicev M, Muglia LJ, Abbot P, et al. Comparing human and macaque placental transcriptomes to disentangle preterm birth pathology from gestational age effects. Placenta. 2016;41:74–82.

62. Enquobahrie DA, Williams MA, Qiu C, Muhie SY, Slentz-Kesler K, Ge Z, Sorenson T. Early pregnancy peripheral blood gene expression and risk of preterm delivery: a nested case control study. BMC Pregnancy Childbirth. 2009;9:56.

63. Mannstadt M, Juppner H, Gardella TJ. Receptors for PTH and PTHrP: their biological importance and functional properties. Am J Phys. 1999;277(5 Pt 2): F665–75.

64. Urist MR. Bone: formation by autoinduction. Science. 1965;150(3698):893–9.

65. Edman K, Furber M, Hemsley P, Johansson C, Pairaudeau G, Petersen J, Stocks M, Tervo A, Ward A, Wells E, et al. The discovery of MMP7 inhibitors exploiting a novel selectivity trigger. ChemMedChem. 2011;6(5):769–73.

66. Munoz-Descalzo S, Hadjantonakis AK, Arias AM. Wnt/ss-catenin signalling and the dynamics of fate decisions in early mouse embryos and embryonic stem (ES) cells. Semin Cell Dev Biol. 2015;47-48:101–9.

67. Cunningham TJ, Kumar S, Yamaguchi TP, Duester G. Wnt8a and Wnt3a cooperate in the axial stem cell niche to promote mammalian body axis extension. Dev Dyn. 2015;244(6):797–807.

68. Lako M, Strachan T, Bullen P, Wilson DI, Robson SC, Lindsay S. Isolation, characterisation and embryonic expression of WNT11, a gene which maps to 11q13.5 and has possible roles in the development of skeleton, kidney and lung. Gene. 1998;219(1–2):101–10.

69. Rozario T, DeSimone DW. The extracellular matrix in development and morphogenesis: a dynamic view. Dev Biol. 2010;341(1):126–40.

70. Soh S-E, Tint MT, Gluckman PD, Godfrey KM, Rifkin-Graboi A, Chan YH, Stünkel W, Holbrook JD, Kwek K, Chong Y-S, et al. Cohort profile: Growing Up in Singapore Towards healthy Outcomes (GUSTO) birth cohort study. Int J Epidemiol. 2014;43(5):1401–9.

71. Pan H, Chen L, Dogra S, Teh AL, Tan JH, Lim YI, Lim YC, Jin S, Lee YK, Ng PY, et al. Measuring the methylome in clinical samples: improved processing of the Infinium Human Methylation450 BeadChip Array. Epigenetics. 2012;7(10):1173–87.

72. Johnson WE, Rabinovic A, Li C. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118–27.

73. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013;8(2):203–9.

74. Price ME, Cotton AM, Lam LL, Farre P, Emberly E, Brown CJ, Robinson WP, Kobor MS. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. Epigenetics Chromatin. 2013;6(1):4.

75. Meng H, Joyce AR, Adkins DE, Basu P, Jia Y, Li G, Sengupta TK, Zedler BK, Murrelle EL, van den Oord EJ. A statistical method for excluding non-variable CpG sites in high-throughput DNA methylation profiling. BMC Bioinformatics. 2010;11:227.

76. Chen J, Just AC, Schwartz J, Hou L, Jafari N, Sun Z, Kocher JP, Baccarelli A, Lin X. CpGFilter: model-based CpG probe filtering with replicates for epigenome-wide association studies. Bioinformatics. 2016;32(3):469–71.

77. Lin X, Tan JYL, Teh AL, Lim IY, Liew SJ, MacIsaac JL, Chong YS, Gluckman PD, Kobor MS, Cheong CY, et al. Cell type-specific DNA methylation in neonatal cord tissue and cord blood: a 850K-reference panel and comparison of cell types. Epigenetics. 2018, in press. https://doi.org/10.1080/15592294.2018.1522929.

78. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics. 2016; 32(2):286–8.

79. Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011;6(7):e21800.

80. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882–3.

81. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, Geman D, Baggerly K, Irizarry RA. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet. 2010;11(10):733–9.