BMC Genomics

The idiopathic preterm delivery methylation profile in umbilical cord blood DNA

Febilla Fernando[1*], Remco Keijser[1], Peter Henneman[2], Anne-Marie F. van der Kevie-Kersemaekers[2], Marcel MAM Mannens[2], Joris AM van der Post[3], Gijs B. Afink[1] and Carrie Ris-Stalpers[1,3]

Abstract

Background: Preterm delivery is the leading cause of neonatal morbidity and mortality. Two-thirds of preterm deliveries are idiopathic. The initiating molecular mechanisms behind spontaneous preterm delivery are unclear. Umbilical cord blood DNA samples are an easy source of material to study the neonatal state at birth. DNA methylation changes can be exploited as markers to identify spontaneous preterm delivery. To identify methylation differences specific to idiopathic preterm delivery, we assessed genome-wide DNA methylation changes in 24 umbilical cord blood samples (UCB) using the 450 K Illumina methylation array. After quality control, conclusions were based on 11 term and 11 idiopathic preterm born neonates. The differentially methylated positions (DMPs) specific for preterm/term delivery, neonatal sex, use of oxytocin and mode of initiation of labor were calculated by controlling the FDR p value at 0.05.

Results: The analysis identifies 1855 statistically significant DMPs between preterm and term deliveries of which 508 DMPs are also attributable to clinical variables other than preterm versus term delivery.
1347 DMPs are unique to term vs preterm delivery, of which 196 DMPs do not relate to gestational age as such. Pathway analysis indicated enrichment of genes involved in calcium signalling, myometrial contraction and relaxation pathways. The 1151 DMPs that correlate with advancing gestational age ($p < 0.05$) include 161 DMPs that match with two previously reported studies on UCB methylation.
Additionally, 123 neonatal sex specific DMPs, 97 DMPs specific to the induction of labour and 42 DMPs specific to the mode of initiation of labor were also identified.

Conclusion: This study identifies 196 DMPs in UCB DNA of neonates which do not relate to gestational age or any other clinical variable recorded and are specific to idiopathic preterm delivery. Furthermore, 161 DMPs from our study overlap with previously reported studies of which a subset is also reported to be differentially methylated at 18 years of age. A DMP on *MYL4*, encoding myosin light chain 4, is a robust candidate for the identification of idiopathic preterm labour as it is identified by all 3 independent studies.

Keywords: Epigenetics, Methylation, UCB, Preterm, Labor, Oxytocin

Background

Preterm birth, defined as delivery before 37 weeks of gestation, has a global prevalence of 9.6 %. It is the leading cause of neonatal morbidity and mortality and is responsible for approximately 70 % of all neonatal deaths and 40 % of childhood neurological morbidities [1, 2]. One third of preterm births are iatrogenic because of maternal or fetal reasons, usually in relation to maternal hypertensive disease. Two thirds occur as a consequence of idiopathic preterm delivery initiated by either spontaneous preterm labor (PtLb) with intact membranes or preterm premature rupture of membranes (PPROM) [3, 4].

The initiating mechanisms behind the inappropriate early activation of labor are poorly understood. Known risk factors include uterine anomalies, multiple gestations, polyhydramnios, blood loss, infection and previous preterm birth [3]. There are no early diagnostic or prognostic markers for spontaneous preterm birth.

* Correspondence: j.f.fernando@amc.uva.nl
[1]Reproductive Biology Laboratory, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands
Full list of author information is available at the end of the article

Fernando *et al. BMC Genomics* (2015) 16:736

Page 2 of 12

The timing of delivery is an intricate play between the fetus, uterus, decidua and placenta. Fetal growth increasing the tension in the uterine wall, increased oxytocin and estrogen bioactivity, progesterone withdrawal and inflammatory decidual activation, all contribute to the initiation of contractions [3, 5]. For several populations susceptibility loci influencing preterm birth have been reported, but until now this has not resulted in the identification of additional general mechanisms involved in preterm labor [6–9]. It is currently not known if there are other fetal properties, apart from fetal growth, that relate to the premature onset of delivery.

Epigenetics is an important regulatory mechanism contributing to the control of gene expression. It has been suggested that epigenetics is involved in idiopathic preterm delivery in humans [10]. DNA methylation in myometrium is known to contribute to the functional progesterone withdrawal associated with labor [11, 12]. The amnion shows significant methylation changes in the promoter region of the oxytocin receptor gene between preterm and term labor [13]. Cervical DNA methylation has also been associated with gestational length [14].

Although a few studies reported the differential DNA methylation profile of umbilical cord blood (UCB) in relation to preterm delivery, the overall emerging picture is complicated by variability in mode of delivery, limitations to specific ethnicity, the absence of information on the progression of labor after initiation of contractions and the inclusion of patients with hypertensive disease [15–17].

In the current study, we investigate the genome-wide methylation profile in UCB from 12 preterm and 12 matched term neonates all born after vaginal delivery from normotensive pregnancies and identify genes that could function as leads to establish early diagnostic markers for idiopathic preterm delivery.

Results
Sample cohort
We selected 24 UCB samples from neonates born after normotensive singleton pregnancies that presented head-first and were vaginally delivered (see Additional file 1: Table S1). Gestational age at delivery ranged from 26 weeks and 3 days to 41 weeks and 4 days; delivery before 37 weeks of gestation was considered preterm. Two samples had to be excluded because of reasons described extensively below resulting in the analysis of 22 samples. Neonatal birth percentile specific for parity, neonatal sex and gestational age was calculated based on the Dutch neonatal growth charts (www.perinatreg.nl). Apart from the gestational age of delivery, neonatal weight and treatment with antenatal glucocorticoids, all characteristics intrinsic to the gestational age at birth, there are no clinical differences between groups (see Table 1).

Neonatal sex chromosomal variations detected by DNA methylation profiling of umbilical cord blood
The Illumina 450 K BeadChip array was used to determine methylation status in our UCB cohort. As an initial check for sample integrity, the 24 samples of the original dataset were grouped as males or females based on their phenotype at birth (male: $n = 12$; female: $n = 12$; Additional file 1: Table S1) Multi-dimensional scaling (MDS) analysis was performed on raw intensity signals as a quality control step using the ChAMP package, which resulted in a male neonate born at term (T4) clustering together with the females in our dataset (Fig. 1a). Levels of probe intensity suggested that this sample gained one X chromosome. This finding was validated on DNA isolated from placenta tissue from this pregnancy that was available through our PANDA biobank project using quantitative fluorescent polymerase chain reaction (QF-PCR) (see Fig. 1b). The QF-PCR signal pattern for markers DXS6803, *HPRT*, and *TAF9* (regions Xq21.31, Xq26.2 and Xq21.1 respectively) are consistent with the presence of 2 X chromosomes. The signal from *SRY* (Yp11.31) marker confirms the presence of a Y chromosome. The signal pattern for *AMEL* (Xp22.22/Yp11.2) is consistent with two X chromosomes and one Y chromosome. The copy number analysis for chromosome X and Y showed evidence of an XXY chromosomal pattern concordant with Klinefelter syndrome. This sample was excluded from further methylation analysis (Additional file 1: Table S1).

Genome-wide methylation differences in umbilical cord blood DNA between preterm and term infants at birth
The array quality metrics package [18] was used to check for quality of the array data and we identified one outlying sample with signal intensity strongly deviating from all other samples. This sample (PT9) was excluded from further analysis (Additional file 1: Table S1).

Differential methylation analysis was performed on 11 preterm and 11 term DNA samples isolated from UCB. We identified 1855 statistically significant differentially methylated positions (DMPs) (FDR $p < 0.05$), of which, 1347 DMPs can be solely attributed to methylation differences between infants born preterm and term. For the comparisons male vs female, no-stimulation vs stimulation with oxytocin and initiation with PtLb vs PPROM, the number of DMPs unique for the comparison groups are 123, 97 and 42 respectively (Fig. 2a, Additional file 2: Table S2, Additional file 4: Table S4a, Additional file 5: Table S5a, Additional file 6). The remaining 246 DMPs were shared by 2 or more groups (Fig. 2a, Additional file 7: Table S7).

For all comparison groups 35-40 % of DMPs are present in gene bodies (Fig. 2b) reflecting the distribution of

Fernando *et al. BMC Genomics* (2015) 16:736

Page 3 of 12

Table 1 Maternal and fetal characteristics of the UCB study cohort

Clinical parameter	Preterm delivery	Term delivery	*P*-value
Number of samples	11	11	1**
Nulliparous	5 (50 %)	6 (55 %)	0.99**
Received antenatal glucocorticoids	9 (82 %)	0 (0 %)	0.0002**
Parturition initiated with spontaneous labor	7 (64 %)	7 (64 %)	1**
Parturition initiated with spontaneous rupture of membranes	4 (36 %)	4 (36 %)	1**
Labor induced or stimulated with oxytocin	5 (45 %)	7 (64 %)	0.67**
Mean gestational age at delivery in weeks + days (range)	29^{+6} (26^{+3} – 36^{+4})	39^{+6} (38^{+0} - 41^{+4})	3.83E-07*
Male neonate	5 (45 %)	6 (55 %)	0.99**
Neonatal weight (g)	1550 (990–2615)	3320 (2525–4215)	7.85E-07*
Neonatal birth weight percentile <10	0 (0 %)	3 (9.09 %)	0.21**
Maternal age	28 (19–35)	28 (19–36)	0.51*

Data are represented as numbers (%) median (range)
Abbreviations: n.s. not significant
P-values were calculated by unpaired student's T-test* and Fisher's exact test** as applicable

probes available on the array, of which 33 % are present in gene bodies.

Correlation of unique term vs preterm DMPs with clinical variables

A hierarchical clustering analysis was performed, in order to identify the clinical variables that associated with the 1347 DMPs identified in the term vs preterm group (Fig. 3a). The DMPs group into 3 main clusters based on gestational age; strictly preterm (26^{+3} till 31^{+3}), intermediate (29^{+6} till 38^{+4}) and term (38^{+3} till 41^{+4}) (Fig. 3b). 1-way ANOVA with 2 degrees of freedom shows that the median gravidity of the intermediate group is significantly higher when compared to the strictly preterm group (adjusted $p = 0.04$). A Pearson correlation test was used to identify the degree of correlation of every DMP with gestational age controlling the FDR at 0.05. Of the total 1347 DMPs, 1151 DMPs correlate significantly with gestational age of which 762 DMPs with a positive and 389 DMPs with a negative correlation to gestational age (Fig. 3c). The DMPs located on gene regions with the highest correlation of gestational age with a negative correlation are *NCOR2*, *DNAJC17*, *PYCR2*, *ATP6V0A1*, *RARA* and *FBLN7* and DMPs with a positive correlation are *IGF2BP1*, *OTOF*, *ATP2B2*, *CES3* and *MYH7B*. The methylation trend of the majority of gestational age related DMPs increases with advancing gestation and within this set are 3 DMPs localising to the *ESR1* gene (see Additional file 2: Table S2, colour shading reflecting Fig. 3c).

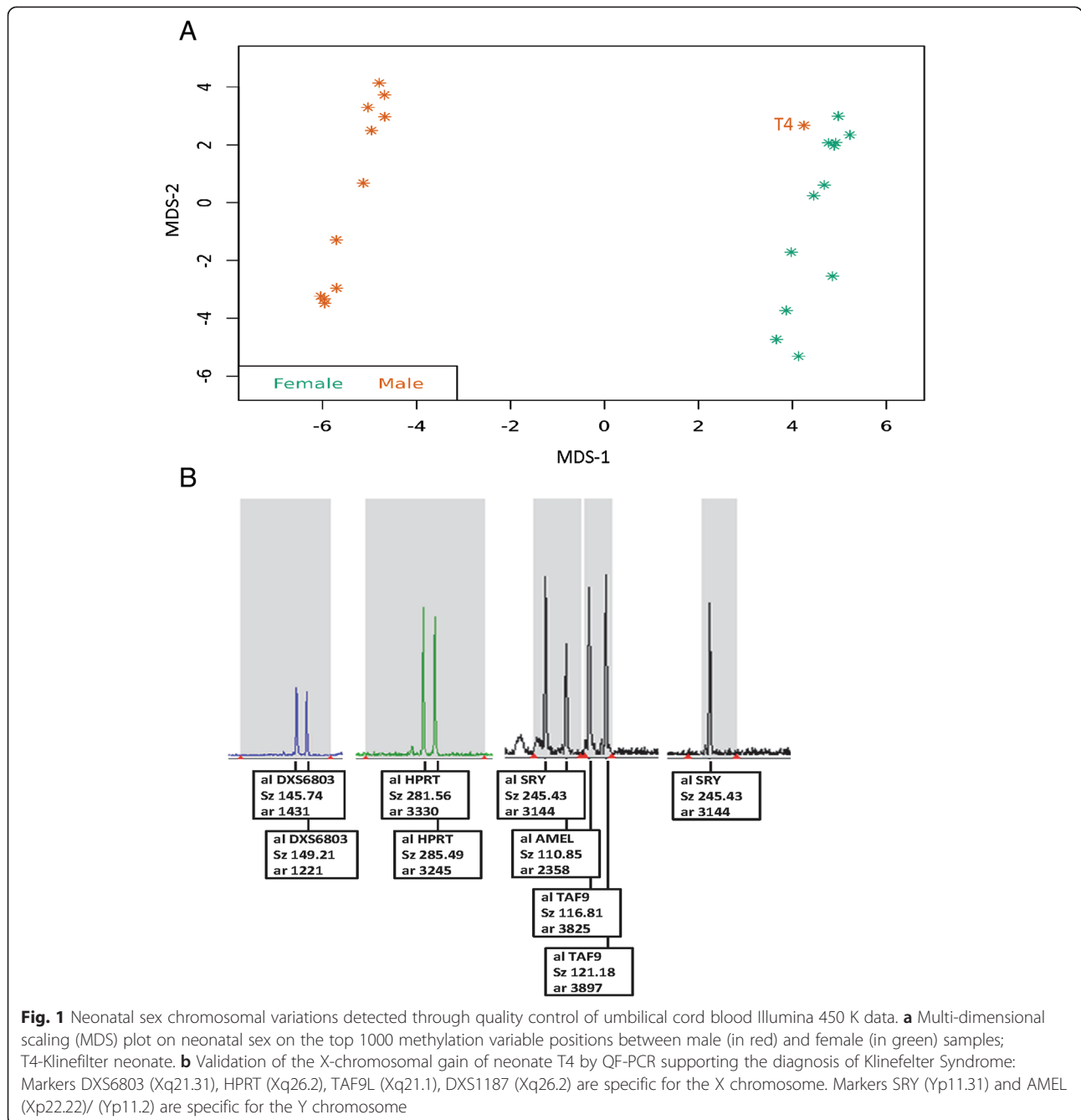The remaining 196 DMPs that do not correlate with gestational age reflect the systematic difference between both groups. DMPs on *KIAA0513*, *UBE21* and *AP3D1* have the highest absolute difference between groups with a p value <0.05 (see Additional file 3: Table S3a).

Pathway analysis

Pathway analysis was performed for all the DMPs identified in each individual comparison group. The gene symbols of the observed DMPs were imported into Web-Gestalt [19, 20] and wiki pathway enrichment analysis was performed with FDR at 0.05. The 42 DMPs specific for the comparison of initiation with PtLb vs PPROM could not be assigned to specific pathways. Among the top enriched pathways in the term vs preterm UCB group, are MAPK signalling, myometrial relaxation and contraction pathway and TGF beta signalling pathway ($p < 0.05$) (Additional file 3: Table S3). The pathway involving calcium regulation in cardiac cell is observed in all 3 comparison groups. Both the term vs preterm and the male vs female comparisons showed an enrichment in myometrial relaxation and contraction pathway, insulin signalling and signalling pathways in glioblastoma. Although the pathways indicated are the same, the genes allocated to them are different, with the exception of *SLC8A1*.

Linkage to other studies

To complement our approach of methylation analysis with relevance to preterm birth we compared DMPs identified in our study to DMPs identified by two other studies on cord blood methylation using the Illumina BeadChip technology (Fig. 4). In total, 161 of the 1347 DMPs that are uniquely assigned to the term versus preterm comparison have been reported previously [15, 17]. We identified DMPs on *ESR1*, *TMEM184A* and *GP1BB* that have been previously associated with gestational age

Fernando *et al. BMC Genomics* (2015) 16:736

Page 4 of 12

Fig. 1 Neonatal sex chromosomal variations detected through quality control of umbilical cord blood Illumina 450 K data. **a** Multi-dimensional scaling (MDS) plot on neonatal sex on the top 1000 methylation variable positions between male (in red) and female (in green) samples; T4-Klinefilter neonate. **b** Validation of the X-chromosomal gain of neonate T4 by QF-PCR supporting the diagnosis of Klinefelter Syndrome: Markers DXS6803 (Xq21.31), HPRT (Xq26.2), TAF9L (Xq21.1), DXS1187 (Xq26.2) are specific for the X chromosome. Markers SRY (Yp11.31) and AMEL (Xp22.22)/ (Yp11.2) are specific for the Y chromosome

in the Schroeder study [15]. DMPs in 157 genes, including those on *IGF2BP1*, *ADORA2A* and *GABBR1*, have been reported by Cruickshank study [17]. A single DMP on the gene encoding myosin light chain 4 (*MYL4*), in the Transcription start site 200 (TSS200) or 5´Untranslated region (5´UTR) depending on the definition of the 5´ prime end of the transcript, is identified by all 3 studies. The current study confirms 32 of the total 109 DMPs that were differentially methylated at birth and

18 years of age originally identified in the Cruikshank study. These 32 DMPs correspond to 17 genes (Fig. 4).

The 196 DMPs in the current UCB study that show no correlation with gestational age do not overlap with the 29 DMPs that associate with preterm birth adjusted for gestational age reported previously. Of these 29, the DMP cg03706951 in our study shows a negative correlation with gestational age. The other 28 are not identified as DMPs in the current study [16].

Fernando *et al. BMC Genomics* (2015) 16:736

Page 5 of 12

Fig. 2 Differentially methylated positions (DMPs) in umbilical cord blood DNA between preterm and term infants. **a** Venn Diagram depicting the differentially methylated CpG sites within each of the four comparison groups (term versus preterm delivery, male versus female neonates, vaginal delivery with contraction stimulated by oxytocin versus not stimulated and whether the process of delivery initiated with PtLb versus PPROM as well as overlap between groups and subgroups). False Discovery Rate (BH: 0.05). **b** Illustration of the distribution of DMPs over gene regions for all comparison groups. The TSS200: region 200 base pairs within the transcription start site (TSS); TSS: region 1500 base pairs within the TSS excluding the TSS 200 region; UTR: untranslated region as present in the mRNA molecule respectively 5′ of the transcription start site (5′UTR) and 3′ of the termination signal (3′UTR); Body: coding and non-coding regions from the TSS until the termination codon; IGR: intergenic region

UCB Differentially Methylated Regions (DMRs) between term and preterm neonates
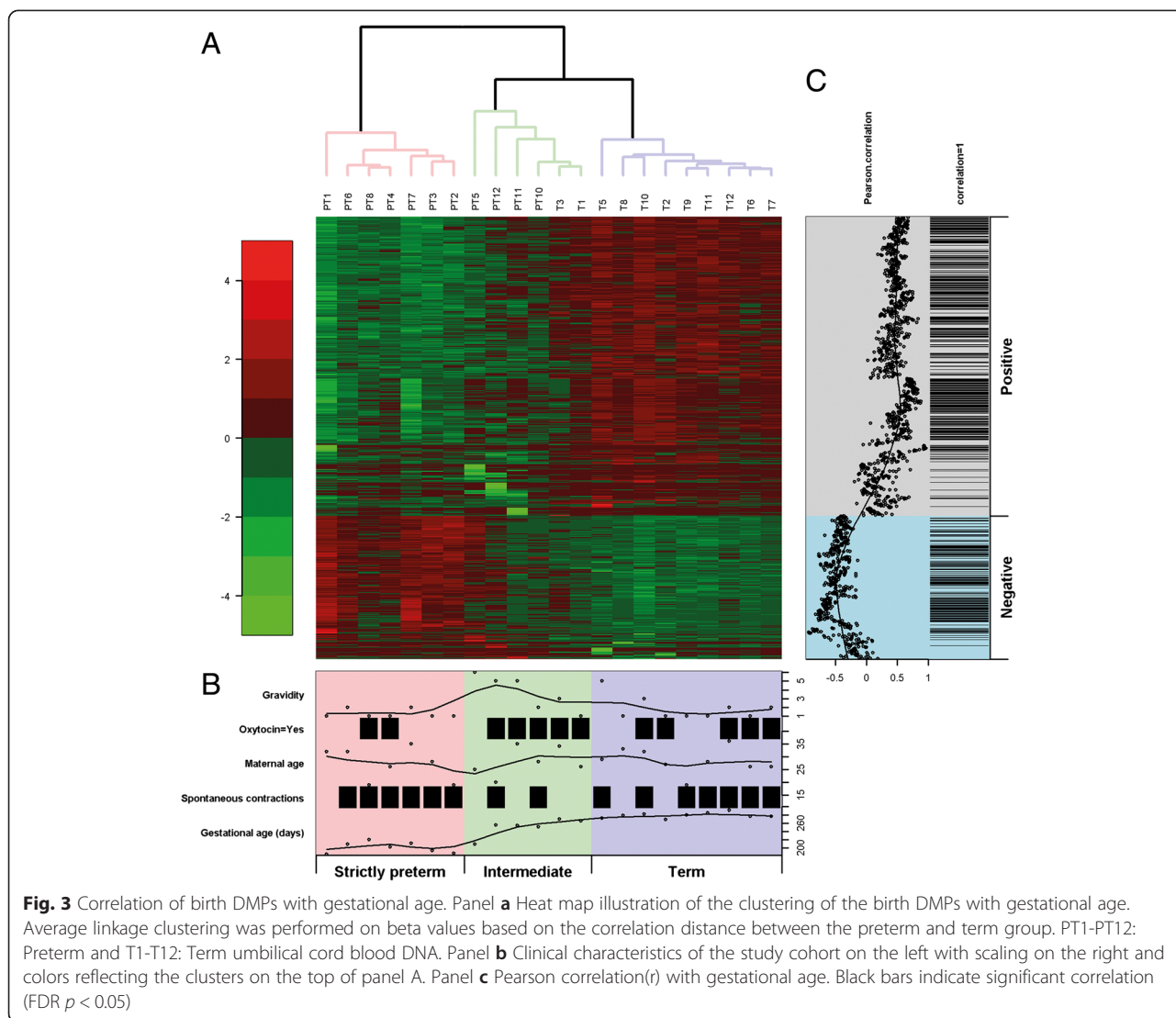
With a FDR at 0.05, we identified 20 differentially methylated regions (DMRs) between preterm and term neonates; 2 non-annotated regions, a non-coding RNA (gene ID: 100131213) and 17 genes (Table 2). Out of the 20, 6 of the identified DMRs are located on chromosome 6. The genes with DMRs were queried for GO biological processes. The top 4 DMRs are located on genes involved in neuronal development such as *PPT2*, *GABBR1*, *PLEKHB1* and induction of immunity such as *ZC3H12D* [21]. Interestingly *ADORA2A*, a gene involved in vascular smooth muscle contraction pathway, was identified as a DMR covering 4 DMPs in this analysis (Fig. 5).

Discussion

DNA methylation plays an important role in gene regulation and in the etiology of complex diseases [22]. UCB

DNA is an easily available source and relevant for the fetal state at the moment of birth. The umbilical cord blood DNA comprises of a mixed population of all types of blood cells and other tissue cells that have entered the circulation. However, the extent to which the DNA methylation profile of UCB reflects those in other tissues is still a matter of debate. Fetal DNA is detectable in the maternal circulation early in pregnancy facilitating non-invasive prenatal testing [23, 24]. Recent developments have shown that the fetal methylome is also accessible from maternal plasma, thus enhancing the opportunity to determine fetal epigenetic markers in the maternal circulation [25].

In the current study, we characterize the methylation profile in UCB DNA specific to spontaneous preterm delivery, neonatal sex, use of oxytocin and mode of initiation of labor using the Illumina 450 K bead chip. This technical approach is also evidently able to reveal

Fernando *et al. BMC Genomics* (2015) 16:736

Page 6 of 12

Fig. 3 Correlation of birth DMPs with gestational age. Panel **a** Heat map illustration of the clustering of the birth DMPs with gestational age. Average linkage clustering was performed on beta values based on the correlation distance between the preterm and term group. PT1-PT12: Preterm and T1-T12: Term umbilical cord blood DNA. Panel **b** Clinical characteristics of the study cohort on the left with scaling on the right and colors reflecting the clusters on the top of panel A. Panel **c** Pearson correlation(r) with gestational age. Black bars indicate significant correlation (FDR *p* < 0.05)

copy number aberrations, as we encountered a 47, XXY neonate in our sample cohort.

There is evidence suggesting gender-specific differences to the susceptibility and progression of some diseases [26–28]. Some studies have shown that preterm born females have a better survival rate and neurological outcome compared to preterm males [29–32]. There is no functional data proving epigenetics as causal to gender differences but the current study indicated some putative associations. We identified 123 autosomal DMPs that differ between males and females irrespective of the preterm/term status. Within this group are DMPs located on *PRKCZ, GBE1 and GYG. PRKCZ* in cord blood T cells has been associated with allergy risk in infants and it has been reported that gender differences affect susceptibility to the development of hypersensitivity reactions [33, 34]. The *PRKCZ* gene has also been linked to development of bipolar disorder [35] in line with the

increased risk of psychiatric illness such as bipolar disorder in males compared to females [36, 37].

Similarly we also established DMPs in certain CpG sites specific to oxytocin usage to induce or stimulate labor. Some of the DMPs in the oxytocin vs no-oxytocin group are *CACNA1D, ADCY9 and ITGA3.* According to KEGG pathways both *CACNA1D* and *ADCY9* are linked to the oxytocin signalling pathway. In-vivo studies show that α3β1 integrin (*ITGA3*) mediated signalling contributes to the control of amplitude and duration of contractions in response to oxytocin in mammary myoepithelial cells [38].

The number of DMPs specific to spontaneous contractions vs spontaneous rupture of membranes comparison is too small to perform pathway analysis.

The majority of methylation differences in our study cohort are between preterm versus term pregnancies. The 1347 DMPs solely attributable to term versus
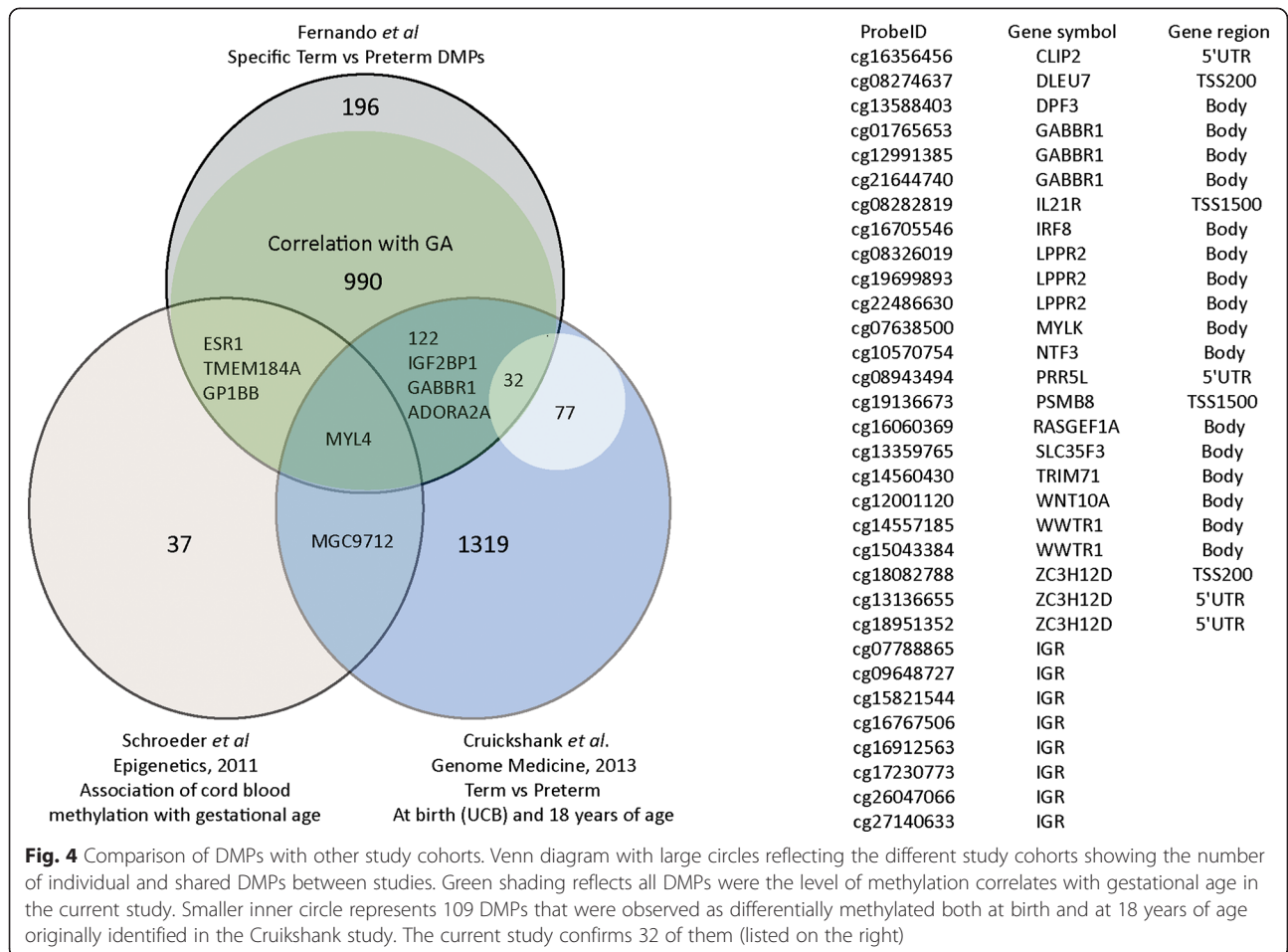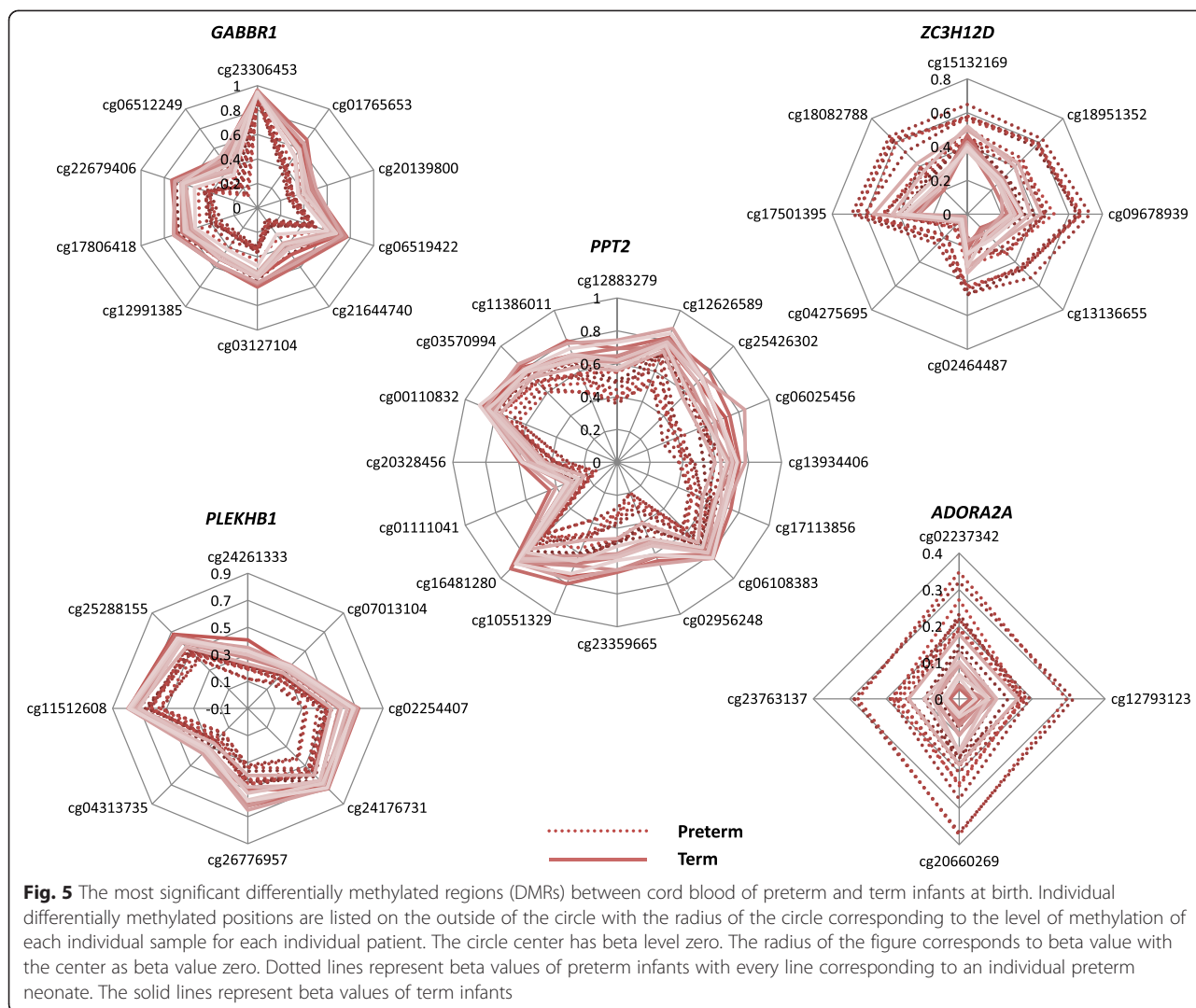
Fernando et al. BMC Genomics (2015) 16:736

Page 7 of 12

ProbeID	Gene symbol	Gene region
cg16356456	CLIP2	5'UTR
cg08274637	DLEU7	TSS200
cg13588403	DPF3	Body
cg01765653	GABBR1	Body
cg12991385	GABBR1	Body
cg21644740	GABBR1	Body
cg08282819	IL21R	TSS1500
cg16705546	IRF8	Body
cg08326019	LPPR2	Body
cg19699893	LPPR2	Body
cg22486630	LPPR2	Body
cg07638500	MYLK	Body
cg10570754	NTF3	Body
cg08943494	PRR5L	5'UTR
cg19136673	PSMB8	TSS1500
cg16060369	RASGEF1A	Body
cg13359765	SLC35F3	Body
cg14560430	TRIM71	Body
cg12001120	WNT10A	Body
cg14557185	WWTR1	Body
cg15043384	WWTR1	Body
cg18082788	ZC3H12D	TSS200
cg13136655	ZC3H12D	5'UTR
cg18951352	ZC3H12D	5'UTR
cg07788865	IGR	
cg09648727	IGR	
cg15821544	IGR	
cg16767506	IGR	
cg16912563	IGR	
cg17230773	IGR	
cg26047066	IGR	
cg27140633	IGR	

Fig. 4 Comparison of DMPs with other study cohorts. Venn diagram with large circles reflecting the different study cohorts showing the number of individual and shared DMPs between studies. Green shading reflects all DMPs were the level of methylation correlates with gestational age in the current study. Smaller inner circle represents 109 DMPs that were observed as differentially methylated both at birth and at 18 years of age originally identified in the Cruikshank study. The current study confirms 32 of them (listed on the right)

Table 2 Gene annotations of differentially methylated regions in UCB DNA from term and preterm born neonates

Gene ID	Gene symbol	Number of DMPs	Adjusted p value	Chromosomal location	DMR start position	DMR end position
9374	PPT2	16	1,78E-12	6p	32120761	32121161
780	DDR1	13	9,46E-06	6p	30853990	30854352
2550	GABBR1	10	4,15E-08	6p	29599012	29599538
58473	PLEKHB1	8	7,08E-09	11q	73356997	73357675
340152	ZC3H12D	8	1,13E-06	6q	149805798	149806410
57224	NHSL1	7	2,21E-07	6q	138866338	138867573
23007	PLCH1	6	3,45E-07	3q	155421691	155422249
3200	HOXA3	5	7,80E-05	7p	27153536	27153782
79899	PRR5L	5	2,08E-07	11p	36422098	36422894
135	ADORA2A	4	1,68E-05	22q	24823473	24823635
3141	HLCS	4	1,46E-05	21q	38362713	38362783
2247	FGF2	3	1,52E-06	4q	123747468	123747623
51199	NIN	3	1,25E-06	14q	51296493	51296731
5696	PSMB8	3	6,23E-07	6p	32812868	32813030
56963	RGMA	3	9,13E-07	15q	93617020	93617110
10418	SPON1	3	1,83E-07	11p	13983766	13983811
202915	TMEM184A	3	1,95E-06	7p	1595968	1596132

Fernando *et al. BMC Genomics* (2015) 16:736

Page 8 of 12

Fig. 5 The most significant differentially methylated regions (DMRs) between cord blood of preterm and term infants at birth. Individual differentially methylated positions are listed on the outside of the circle with the radius of the circle corresponding to the level of methylation of each individual sample for each individual patient. The circle center has beta level zero. The radius of the figure corresponds to beta value with the center as beta value zero. Dotted lines represent beta values of preterm infants with every line corresponding to an individual preterm neonate. The solid lines represent beta values of term infants

preterm birth were mainly observed in gene bodies, intergenic regions and within 1500 basepairs of the transcription start site, reflecting the number of probes present in these regions. About 85 % of the 1347 DMPs observed in UCB of neonates born term vs preterm have a clear relation with advancing gestational age. This has been reported before and is assumed to reflect the haematological changes as leucocyte content, nucleated reticulocyte content and maturation of red blood cells, all of which strongly correlate with gestational age [39–42]. In the current cohort gestational ages range from 26^{+3} to 41^{+4} weeks. Hierarchical clustering of the 1347 DMPs separates early preterm from term deliveries and identifies an intermediate gestational age group (29^{+6} – 38^{+4}) with a statistically significant higher gravidity. The remaining 196 DMPs represent which do not correlate with gestational age represent systematic

differences between the preterm and term born neonates and have not been reported previously in relation to preterm delivery. A study interrogating methylation patterns in UCB DNA from African-Americans identified 29 CpG sites associated with preterm birth adjusted for gestational age. None of those were corroborated by the current study [16]. Wiki pathway analysis and gene enrichment of our set of 196 DMPs shows calcium and TGF beta signalling, myometrial contraction and relaxation pathways and the corticotrophin releasing hormone pathway. It is well established that calcium signalling is crucial in the human myometrium for initiation of labor [43]. The genes which are represented in the myometrial contraction and relaxation pathways are sodium and calcium exchanger *SLC8A1*, which also plays a role in cardiomyocyte contractions, chemokine receptor *CXCR7*, guanine nucleotide binding protein *GNG7*

Fernando *et al. BMC Genomics* (2015) 16:736

Page 9 of 12

and transmembrane signalling enzyme phospholipase C *PLCG2* respectively. Further research will have to determine if they are putative foetal biomarkers for idiopathic preterm delivery that can be analysed from the maternal circulation.

Ideally high-thoughput data from clinical samples need to be validated either by additional experiments or comparative validation to other studies. The limitation of the current study is that we have not experimentally validated this finding. However, we have approached this issue with detailed comparison of previously published independent studies on UCB methylation. There have been 3 previous reports on DNA methylation profiles in UCB analysed with respect to gestational age at delivery using the Illumina platform [15–17]. The study by Schroeder et al. analysed 453 UCB samples with gestational age ranging from 32 to 43 weeks of gestation on a 27 K platform. Results were adjusted for a number of clinical variables including neonatal sex and parity and they showed that associations were independent of the method of delivery or induction of labor. The study by Cruickshank et al., interrogated both UCB and blood at 18 years of age of 12 preterm (25 to 30 weeks of gestation) and 12 term (36–42 weeks of gestation) born neonates using the 450 K platform. Two thirds of the mothers did not experience labor. For both these studies gravidity was not reported.

The 161 DMPs which overlap between our study and the above mentioned studies provide increased evidence of presence of these methylation differences in UCB of preterm and term infants at birth. The overlapping DMPs on *ESR1* and *MYL4* are of interest in the context of preterm delivery. Increased *ESR1* (estrogen receptor) gene expression in the myometrium is triggered by functional progesterone withdrawal and results in increased estrogen bioactivity, an important contributor to the transition from myometrial quiescence to synchronised contractions [44, 45]. As all our neonates were delivered vaginally, we cannot exclude the fact that labor is associated with the differential methylation of the *ESR1* gene. However, based on the study of Schroeder which also included samples from neonates delivered by caesarean section, this seems unlikely.

MYL4 encodes myosin light chain 4 that is essential to the myometrial contraction pathway. *MYL4* expression in mice myometrium is downregulated over two-fold during quiescence indicating an active role in the progression to myometrial contractions [46].

A single DMP (cg19817652) on *C17orf98* identified by the Cruickshank study as relating to gestational age is uniquely attributable to neonatal sex in our study. The 32 overlapping DMPs with Cruickshank study which has been experimentally validated to be present at 18 years of age is of significance due to the long-term consequence of preterm birth.

DMR analysis identifies 20 significant differentially methylated regions. The most significant DMR is located on the *PPT2* gene on chromosome 6p21.3 and covers 16 DMPs. *PPT2* is located on the major histocompatibility locus (MHC-III) and its deficiency in homozygous knock out mice causes a neurodegenerative disorder [47]. There are 2 other DMRs present on 6p21.3 corresponding to *DDR1* with 13DMPs and *PSMB8* with 3DMPs. *ZC3H12D* activates TLR signalling in macrophages and plays a role in immunity and inflammatory diseases. The inflammatory pathway is a well-established important contributor to the initiation of synchronised myometrial contractions and interestingly the differential methylation of *ZC3H12D* has been shown to persist until at least the age of 18 years [48]. The *ADORA2A* encodes the adenosine A2a receptor, a member of the G-protein coupled receptor superfamily. Interleukin-1b activates myometrial inflammation and is able to up regulate *ADORA2A* expression >20 fold in primary culture of pregnant human uterine myocytes [49]. *ADORA2A*, also reported by the Cruickshank study, shows a relative increased methylation status in UCB DNA of neonates born after idiopathic preterm labor.

Comparing the methylation and gene expression profiles of myometrium, placenta and umbilical cord blood of preterm and term deliveries could shed more light on the functional relevance of the methylation changes observed in the cord blood DNA.

Conclusion

This study identifies 1347 methylation changes in umbilical cord blood specific to idiopathic preterm delivery. The approach to methylation analysis in this study also delineates methylation changes specific to neonatal sex giving more insight into sex-specific autosomal methylation differences at birth enhancing the understanding of gender specific susceptibility to a specific disease.

In addition, the identification of 161 DMPs which overlap with other studies offer supportive evidence of robust methylation differences in UCB between preterm and term infants. Meta-analysis of raw data sets of these independent studies combined with detailed information on clinical data will provide a more robust set of DMPs which can be used as diagnostic markers for spontaneous preterm delivery.

Methods
Study design
From the Preeclampsia And Non-preeclampsia Database (PANDA) project we used placenta tissue and umbilical cord genomic DNA with clinical data on maternal health, the course of pregnancy and pregnancy outcome with informed consent, approved by the ethics committee of the Academic Medical Center Hospital of the University of the Amsterdam.

Fernando *et al. BMC Genomics* (2015) 16:736

Page 10 of 12

We selected 24 umbilical cord blood (UCB) samples from vaginal deliveries of normotensive singleton pregnancies where the live born neonate presented head first and the membranes ruptured spontaneously, either before or after the start of contractions. Pregnancies with maternal diabetes, hyper- or hypothyroidism, preeclampsia, HELLP syndrome, congenital heart disease, or intrahepatic cholestasis of pregnancy, intra-uterine growth retardation, fetal distress or any kind of infection were excluded. We selected UCB from 12 neonates delivered <37 weeks of gestation and they were matched to 12 neonates delivered ≥ 37 weeks of gestation with respect to parity, maternal age, use of oxytocin, initiation of parturition with spontaneous labor or rupture of membranes and neonates small for gestational age. UCB was collected in EDTA tubes and processed by Gentra Autopure LS98™ system (Gentra Systems). Table 1 illustrates the maternal and fetal characteristics of the cord blood study cohort (Additional clinical data in Additional file 1: Table S1).

Sample preparation and methylation assay

For each sample, 1ug of genomic DNA was bisulphite converted using the EZ-DNA methylation kit (Zymo Research, D5001). The bisulphite converted DNA was subjected to whole genome amplification, fragmentation followed by hybridization on Infinium Illumina 450 K BeadChip array and Cy3 or Cy5 signal intensities were generated using Illumina iScan.

Quantitative fluorescence-PCR analysis

DNA was isolated from a placenta biopsy using the MagnaPureLC DNA isolation kit II (Roche). For QF-PCR aneuploidy screening the QST*Rplus V2 kit (Elucigene Diagnostics, AN0PLB2) was used according to the manufacturer's instructions. The amplified sample was analyzed with the ABI 3500 (Applied Biosystems, Foster City, California, USA). Interpretation of results was performed using guidelines from the manual, the 2012 ACC/CMGS 'QF-PCR for the diagnosis of aneuploidy best practice guidelines' V3.01, 22 and the CCMG 'Practice Guidelines for Prenatal QF-PCR'.

Initial data analysis

The array quality metrics package was used to check for outliers based on array quality that were subsequently excluded from further analysis [18]. Raw data files containing the methylation intensities were imported into the ChAMP package [50]. Additional quality control was performed using ChAMP, and probes with a detection p-value less than $p < 0.01$ were filtered out and used for the analysis scores. The X and Y chromosome methylation levels in UCB were checked against gender of the neonate before they were removed from the analysis altogether.

One sample from the preterm group (PT9) was identified as an outlier based on the distance between arrays and intensity distributions. This sample was excluded from the dataset before normalization. The Beta-mixture quantile normalization (BMIQ) method was used to normalize the dataset and to correct for the Illumina assay-II signal bias within the Illumina 450 K BeadChip [50]. To avoid methylation bias due to cross hybridizing probes, 29,233 probes that were previously reported to cross hybridize were removed from the dataset [51]. After the quality control steps, differential methylation analysis was performed with the remaining 443,190 probes.

Statistical testing

Linear regression modelling algorithms from the LIMMA package [52] were used to construct linear models for each comparison group which include neonatal sex, type of parturition initiation and induction or co-stimulation with oxytocin. Samples were assigned into 4 comparison groups based on 1) preterm or term status 2) neonatal sex: male or female 3) whether oxytocin was administered at any time before delivery 4) whether parturition initiated with PtLb or PPROM. Models were constructed by setting the M value (log transformed beta value) as outcome and the attributed group phenotype as independent variable for each group. The decideTest function was used for multiple testing across all groups. The differentially methylated positions (DMPs) were called based on the p value generated by multiple testing after a false discovery rate correction (FDR) according to Benjamini and Hochberg $p < 0.05$. Gene region information was obtained from the Illumina manifest file version 1.2. After identifying the 1347 DMPs that were exclusive to the preterm vs. term comparison group they were further used to define differentially methylated regions (DMRs) and perform pathway analysis.

Differentially methylated regions

DMRs were calculated using the probe lasso function of the ChAMP package [50]. A gene region was called DMR if it had a minimum of 3 or more significant probes after adjusting for FDR ($p < 0.05$) and a neighbouring DMR is separated by a distance of maximum 1000 base pairs.

Pathway analysis

The DMPs for each comparison group were narrowed down to their gene names and imported into WebGestalt for wiki-pathway analysis [19, 20] A hypergeometric statistical testing was performed and a p value was generated after FDR correction (Benjamini and Hochberg) $p = 0.05$.

Fernando et al. BMC Genomics (2015) 16:736

Page 11 of 12

Availability of supporting data

The methylation data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus [53] and are accessible through GEO Series accession number GSE66459 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66459).

Additional files

Additional file 1: Patient data. (XLS 104 kb)

Additional file 2: DMPs correlating to gestational age (refers to Fig. 3c). (XLSX 140 kb)

Additional file 3: a. Systematic differences (DMPs) Term vs Preterm. b. Term vs Preterm-wikipathways. (ZIP 99 kb)

Additional file 4: a. Systematic differences (DMPs) Female vs Male neonates. b. Females vs Males- wikipathways. (ZIP 30 kb)

Additional file 5: a. Systematic differences (DMPs) Oxcytocin use. b. Oxcytocin use-wikipathways. (ZIP 25 kb)

Additional file 6: Systematic differences (DMPs) Initiation with spontaneous PtLb vs PPROM. (XLS 44 kb)

Additional file 7: DMPs shared between groups. (XLS 155 kb)

Abbreviations

DMP: Differential methylated position; DMR: Differentially methylated region; FDR: False discovery rate; GO: Gene ontology; IGR: Intragenic region; MDS: Multiple dimensional scaling; QF-PCR: Quantitative fluorescent polymerase chain reaction; TSS1500: Within 1500 basepairs of the transcription start site; TSS200: Within 200 basepairs of the transcription start site; UCB: Umbilical cord blood; UTR: Untranslated region.

Competing interests

The authors have no potential conflicts of interest with respect to the research, authorship, and/or publication of this work.

Authors' contributions

JAMvdP/MMAMM/CRS/GBA designed the study. FF collected samples for the study, performed methylation data analysis and interpretation. RK provided critical input in data analysis by R programming. PH helped with preliminary data analyses. AMFvdKK performed the QF-PCR validation for the study. CRS coordinated the study. FF/CRS drafted the manuscript with input and reviews from GBA, RK, PH and AMFvdKK. All authors read and approved the final manuscript.

Author details

[1]Reproductive Biology Laboratory, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.
[2]Department of Clinical Genetics, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.
[3]Women's and Children's Clinic, Department of Obstetrics and Gynaecology, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.

References

1. Hack M, Fanaroff AA. Outcomes of children of extremely low birthweight and gestational age in the 1990s. Semin Neonatol. 2000;5(2):89–106. doi:10.1053/siny.1999.0001.
2. Saigal S, Doyle LW. An overview of mortality and sequelae of preterm birth from infancy to adulthood. Lancet. 2008;371(9608):261–9. doi:10.1016/s0140-6736(08)60136-1.
3. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. Lancet. 2008;371(9606):75–84. doi:10.1016/s0140-6736(08)60074-4.
4. Romero R, Dey SK, Fisher SJ. Preterm labor: one syndrome, many causes. Science. 2014;345(6198):760–5. doi:10.1126/science.1251816.
5. Charpigny G, Leroy MJ, Breuiller-Fouche M, Tanfin Z, Mhaouty-Kodja S, Robin P, et al. A functional genomic study to identify differential gene expression in the preterm and term human myometrium. Biol Reprod. 2003;68(6):2289–96. doi:10.1095/biolreprod.102.013763.
6. Chittoor G, Farook VS, Puppala S, Fowler SP, Schneider J, Dyer TD, et al. Localization of a major susceptibility locus influencing preterm birth. Mol Hum Reprod. 2013;19(10):687–96. doi:10.1093/molehr/gat036.
7. Hao K, Wang X, Niu T, Xu X, Li A, Chang W, et al. A candidate gene association study on preterm delivery: application of high-throughput genotyping technology and advanced statistical methods. Hum Mol Genet. 2004;13(7):683–91. doi:10.1093/hmg/ddh091.
8. Karjalainen MK, Huusko JM, Tuohimaa A, Luukkonen A, Haataja R, Hallman M. A study of collectin genes in spontaneous preterm birth reveals an association with a common surfactant protein D gene polymorphism. Pediatr Res. 2012;71(1):93–9. doi:10.1038/pr.2011.2.
9. Haataja R, Karjalainen MK, Luukkonen A, Teramo K, Puttonen H, Ojaniemi M, et al. Mapping a new spontaneous preterm birth susceptibility gene, IGF1R, using linkage, haplotype sharing, and association analysis. PLoS Genet. 2011;7(2), e1001293. doi:10.1371/journal.pgen.1001293.
10. Boyd HA, Poulsen G, Wohlfahrt J, Murray JC, Feenstra B, Melbye M. Maternal contributions to preterm delivery. Am J Epidemiol. 2009;170(11):1358–64. doi:10.1093/aje/kwp324.
11. Li X, Chen C, Luo H, van Velkinburgh JC, Ni B, Chang Q. Decreased DNA Methylations at the Progesterone Receptor Promoter A Induce Functional Progesterone Withdrawal in Human Parturition. Reprod Sci. 2014;21(7):898–905. doi:10.1177/1933719113518982.
12. 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. doi:10.1095/biolreprod.113.113209.
13. Kim J, Pitlick MM, Christine PJ, Schaefer AR, Saleme C, Comas B, et al. Genome-wide analysis of DNA methylation in human amnion. Sci World J. 2013;2013:678156. doi:10.1155/2013/678156.
14. Burris HH, Baccarelli AA, Motta V, Byun HM, Just AC, Mercado-Garcia A, et al. Association between length of gestation and cervical DNA methylation of PTGER2 and LINE 1-HS. Epigenetics. 2014;9(8):1083–91.
15. Schroeder JW, Conneely KN, Cubells JC, Kilaru V, Newport DJ, Knight BT, et al. Neonatal DNA methylation patterns associate with gestational age. Epigenetics. 2011;6(12):1498–504. doi:10.4161/epi.6.12.18296.
16. Parets SE, Conneely KN, Kilaru V, Fortunato SJ, Syed TA, Saade G, et al. Fetal DNA methylation associates with early spontaneous preterm birth and gestational age. PLoS One. 2013;8(6), e67489. doi:10.1371/journal.pone.0067489.
17. Cruickshank MN, Oshlack A, Theda C, Davis PG, Martino D, Sheehan P, et al. 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. doi:10.1186/gm500.
18. Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics–a bioconductor package for quality assessment of microarray data. Bioinformatics. 2009;25(3):415–6. doi:10.1093/bioinformatics/btn647.
19. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res. 2005;33(Web Server issue):W741–8. doi:10.1093/nar/gki475.
20. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. Nucleic Acids Res. 2013;41.
21. Huang S, Qi D, Liang J, Miao R, Minagawa K, Quinn T, et al. The putative tumor suppressor Zc3h12d modulates toll-like receptor signaling in macrophages. Cell Signal. 2012;24(2):569–76. doi:10.1016/j.cellsig.2011.10.011.
22. Liyanage VR, Jarmasz JS, Murugeshan N, Del Bigio MR, Rastegar M, Davie JR. DNA modifications: function and applications in normal and disease States. Biology. 2014;3(4):670–723. doi:10.3390/biology3040670.

Fernando *et al. BMC Genomics* (2015) 16:736

Page 12 of 12

23. Taglauer ES, Wilkins-Haug L, Bianchi DW. Review: cell-free fetal DNA in the maternal circulation as an indication of placental health and disease. Placenta. 2014;35(Suppl):S64–8. doi:10.1016/j.placenta.2013.11.014.

24. Hudecova I, Sahota D, Heung MM, Jin Y, Lee WS, Leung TY, et al. Maternal plasma fetal DNA fractions in pregnancies with low and high risks for fetal chromosomal aneuploidies. PLoS One. 2014;9(2), e88484. doi:10.1371/journal.pone.0088484.

25. Lun FM, Chiu RW, Sun K, Leung TY, Jiang P, Chan KC, et al. Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA. Clin Chem. 2013;59(11):1583–94. doi:10.1373/clinchem.2013.212274.

26. Ordovas JM. Gender, a significant factor in the cross talk between genes, environment, and health. Gend Med. 2007;4 Suppl B:S111–22.

27. Butts CL, Sternberg EM. Neuroendocrine factors alter host defense by modulating immune function. Cell Immunol. 2008;252(1–2):7–15. doi:10.1016/j.cellimm.2007.09.009.

28. Grootenboer N, Bosch JL, Hendriks JM, van Sambeek MR. Epidemiology, aetiology, risk of rupture and treatment of abdominal aortic aneurysms: does sex matter? Eur J Vasc Endovasc Surg. 2009;38(3):278–84. doi:10.1016/j.ejvs.2009.05.004.

29. Zisk JL, Genen LH, Kirkby S, Webb D, Greenspan J, Dysart K. Do premature female infants really do better than their male counterparts? Am J Perinatol. 2011;28(3):241–6. doi:10.1055/s-0030-1268239.

30. Kent AL, Wright IM, Abdel-Latif ME. Mortality and adverse neurologic outcomes are greater in preterm male infants. Pediatrics. 2012;129(1):124–31. doi:10.1542/peds.2011-1578.

31. Zeitlin J, Saurel-Cubizolles MJ, De Mouzon J, Rivera L, Ancel PY, Blondel B, et al. Fetal sex and preterm birth: are males at greater risk? Hum Reprod. 2002;17(10):2762–8.

32. Zeitlin J, Ancel PY, Larroque B, Kaminski M. Fetal sex and indicated very preterm birth: results of the EPIPAGE study. Am J Obstet Gynecol. 2004;190(5):1322–5. doi:10.1016/j.ajog.2003.10.703.

33. D'Vaz N, Ma Y, Dunstan JA, Lee-Pullen TF, Hii C, Meldrum S, et al. Neonatal protein kinase C zeta expression determines the neonatal T-Cell cytokine phenotype and predicts the development and severity of infant allergic disease. Allergy. 2012;67(12):1511–8. doi:10.1111/all.12027.

34. Chen W, Mempel M, Schober W, Behrendt H, Ring J. Gender difference, sex hormones, and immediate type hypersensitivity reactions. Allergy. 2008;63(11):1418–27. doi:10.1111/j.1398-9995.2008.01880.x.

35. Kandaswamy R, McQuillin A, Curtis D, Gurling H. Tests of linkage and allelic association between markers in the 1p36 PRKCZ (protein kinase C zeta) gene region and bipolar affective disorder. Am J Med Genet B Neuropsychiatr Genet. 2012;159b(2):201–9. doi:10.1002/ajmg.b.32014.

36. Nosarti C, Reichenberg A, Murray RM, Cnattingius S, Lambe MP, Yin L, et al. Preterm birth and psychiatric disorders in young adult life. Arch Gen Psychiatry. 2012;69(6):E1–8. doi:10.1001/archgenpsychiatry.2011.1374.

37. Peacock JL, Marston L, Marlow N, Calvert SA, Greenough A. Neonatal and infant outcome in boys and girls born very prematurely. Pediatr Res. 2012;71(3):305–10. doi:10.1038/pr.2011.50.

38. Raymond K, Cagnet S, Kreft M, Janssen H, Sonnenberg A, Glukhova MA. Control of mammary myoepithelial cell contractile function by α3β1 integrin signalling. EMBO J. 2011;30(10):1896–906. doi:10.1038/emboj.2011.113.

39. Maconi M, Formisano D, Cavalca L, Rolfo A, Cardaropoli S, Danise P. Reticulocyte count and reticulocyte maturation profile in human umbilical cord blood from healthy newborns. Lab Hematol. 2010;16(1):3–7. doi:10.1532/lh96.09012.

40. Bocker MT, Hellwig I, Breiling A, Eckstein V, Ho AD, Lyko F. Genome-wide promoter DNA methylation dynamics of human hematopoietic progenitor cells during differentiation and aging. Blood. 2011;117(19):e182–9. doi:10.1182/blood-2011-01-331926.

41. Walker JC, Smolders MA, Gemen EF, Antonius TA, Leuvenink J, de Vries E. Development of lymphocyte subpopulations in preterm infants. Scand J Immunol. 2011;73(1):53–8. doi:10.1111/j.1365-3083.2010.02473.x.

42. Correa-Rocha R, Perez A, Lorente R, Ferrando-Martinez S, Leal M, Gurbindo D, et al. Preterm neonates show marked leukopenia and lymphopenia that are associated with increased regulatory T-cell values and diminished IL-7. Pediatr Res. 2012;71(5):590–7. doi:10.1038/pr.2012.6.

43. You X, Gao L, Liu J, Xu C, Liu C, Li Y, et al. CRH activation of different signaling pathways results in differential calcium signaling in human pregnant myometrium before and during labor. J Clin Endocrinol Metab. 2012;97(10):E1851–61. doi:10.1210/jc.2011-3383.

44. Smith R, Mesiano S, McGrath S. Hormone trajectories leading to human birth. Regul Pept. 2002;108(2–3):159–64.

45. Hirota Y, Cha J, Dey SK. Revisiting reproduction: prematurity and the puzzle of progesterone resistance. Nat Med. 2010;16(5):529–31. doi:10.1038/nm0510-529.

46. Salomonis N, Cotte N, Zambon AC, Pollard KS, Vranizan K, Doniger SW, et al. Identifying genetic networks underlying myometrial transition to labor. Genome Biol. 2005;6(2):R12. doi:10.1186/gb-2005-6-2-r12.

47. Gupta P, Soyombo AA, Shelton JM, Wilkofsky IG, Wisniewski KE, Richardson JA, et al. Disruption of PPT2 in mice causes an unusual lysosomal storage disorder with neurovisceral features. Proc Natl Acad Sci U S A. 2003;100(21):12325–30. doi:10.1073/pnas.2033229100.

48. Renthal NE, Chen CC, Williams KC, Gerard RD, Prange-Kiel J, Mendelson CR. miR-200 family and targets, ZEB1 and ZEB2, modulate uterine quiescence and contractility during pregnancy and labor. Proc Natl Acad Sci U S A. 2010;107(48):20828–33. doi:10.1073/pnas.1008301107.

49. Lee Y, Sooranna SR, Terzidou V, Christian M, Brosens J, Huhtinen K, et al. Interactions between inflammatory signals and the progesterone receptor in regulating gene expression in pregnant human uterine myocytes. J Cell Mol Med. 2012;16(10):2487–503. doi:10.1111/j.1582-4934.2012.01567.x.

50. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, et al. ChAMP: 450k chip analysis methylation pipeline. Bioinformatics. 2014;30(3):428–30. doi:10.1093/bioinformatics/btt684.

51. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013;8(2):203–9. doi:10.4161/epi.23470.

52. Gordon K. Smyth. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W, editors. Bioinformatics and computational biology solutions using R and bioconductor. 2005. p. 397–420.

53. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30(1):207–10.