Microarray-based analysis of methylation of 1st trimester trisomic placentas from down syndrome, edwards syndrome and patau syndrome

Hatt, L.; M. Aagaard, Mads; Bach, C; Graakjaer, J.; Sommer, S.; Agerholm, I E; Kølvraa, S.; Bojesen, A.

Published in:
PLoS ONE

DOI:
10.1371/journal.pone.0160319

Publication date:
2016

Document version
Også kaldet Forlagets PDF

Document license
CC BY

Citation for published version (APA):
Microarray-Based Analysis of Methylation of 1st Trimester Trisomic Placentas from Down Syndrome, Edwards Syndrome and Patau Syndrome

Abstract

Methylation-based non-invasive prenatal testing of fetal aneuploidies is an alternative method that could possibly improve fetal aneuploidy diagnosis, especially for trisomy 13 (T13) and trisomy 18 (T18). Our aim was to study the methylation landscape in placenta DNA from trisomy 13, 18 and 21 pregnancies in an attempt to find trisomy-specific methylation differences better suited for non-invasive prenatal diagnosis. We have conducted high-resolution methylation specific bead chip microarray analyses assessing more than 450,000 CpGs analyzing placentas from 12 T21 pregnancies, 12 T18 pregnancies and 6 T13 pregnancies. We have compared the methylation landscape of the trisomic placentas to the methylation landscape from normal placental DNA and to maternal blood cell DNA. Comparing trisomic placentas to normal placentas we identified 217 and 219 differentially methylated CpGs for CVS T18 and CVS T13, respectively (delta β > 0.2, FDR < 0.05), but only three differentially methylated CpGs for T21. However, the methylation differences was only modest (delta β < 0.4), making them less suitable as diagnostic markers. Gene ontology enrichment analysis revealed that the gene set connected to the T18 differentially methylated CpGs was highly enriched for GO terms related to “DNA binding” and “transcription factor binding” coupled to the RNA polymerase II transcription. In the gene set connected to the T13 differentially methylated CpGs we found no significant enrichments.

Introduction

Achieving non-invasive prenatal diagnosis of fetal cases of the common chromosome aneuploidies based on circulating free fetal DNA (cffDNA) in maternal plasma has for more than a decade been the goal of many research groups. The overall strategy for this research was to
achieve sufficiently precise quantitation of e.g. chromosome 21 DNA fragments compared to fragments from a reference chromosome, so that a significant rise in chromosome 21 DNA fragments in maternal plasma could be demonstrated in pregnancies carrying a fetus with trisomy 21 (T21).

One approach to accomplish this is to measure circulating free DNA (cfDNA) in maternal blood with sufficient precision to demonstrate a significantly increased amount of e.g. chromosome 21 derived fragments, even when measuring the substantial background of maternal DNA fragments. This strategy has turned out to be successful for T21 when the quantitation is performed by Next generation sequencing (NGS). However, even though NGS has proven effective for T21 testing there are still problems with reliable demonstration of trisomy 13 (T13) and trisomy 18 (T18) with this approach[1]. Furthermore, the ability to detect increased chromosome dosage is very dependent on cfDNA fraction resulting in a minor fraction of non-reportable results.

An alternative approach to improve the method could be to obtain fetal DNA specificity by preferentially targeting fetal DNA sequences in the quantitative analysis. Thereby, likely lowering the limitation coming from the sensibility to low cfDNA fraction. This has been tried by several groups utilizing features such as DNA fragment length or epigenetic signatures, to distinguish the firstly derived DNA from the huge excess of maternal DNA [2–6]. In this context, we have in a previous study, using full genome methylation arrays, compared the methylation status of DNA from 12 placenta samples to DNA from 10 maternal blood samples, all from pregnancies in the first trimester [7]. We identified highly placenta specific epigenetic markers located on chromosome 13, 18, and 21 as well as highly placenta specific epigenetic markers in the regions of several microdeletion syndromes.

An alternative way to improve the discrimination between fetal and maternal cfDNA for non-invasive prenatal testing (NIPT) of the fetal aneuploidies would be to define fetal markers that are not only placenta specific but specific for trisomic placental DNA, preferable in the form of methylation differences. In agreement with this theory, two recent studies have looked into general epigenetic patterns in placental DNA from T21 cases and observed a general hypermethylation in T21 placentas compared to normal placentas [8,9]. We therefore extended our previous study by investigating the methylation landscape in placental DNA from the three common aneuploidies T21, T13, and T18 and compared it to the methylation landscape in DNA from normal placentas and from maternal blood cell (MBC) DNA in an attempt to demonstrate possible methylation differences better suited for NIPT especially for T13 and T18.

We have in addition looked into possible biological relevance of sites that differed in methylation between DNA from normal and trisomic placenta samples, firstly by characterizing the location of such sites within or between CpG islands and secondly by looking at the biological function of genes closest to the candidate sites using gene ontology.

Materials and Methods
Clinical samples
All the samples for the microarray study were sampled from 1st trimester pregnant women, who underwent chorionic villus sampling (CVS) due to increased risk of T21, T18 or T13, estimated by a combination of the nuchal translucency testing and the double test (measuring the plasma protein markers; Pregnancy Associated Plasma Protein A (PAPP-A) and Chorionic Gonadotropin Beta (free β-hCG)). We used 12 CVS samples from T21 pregnancies, 12 CVS samples from T18 pregnancies and 6 CVS samples from T13 pregnancies, all verified by chromosome analysis on CVS samples. We further used the data from our recent study [7] where
we analyzed 12 CVS samples from normal pregnancies and 10 blood samples, from 1st trimester pregnant women with a normal fetus, judged by a normal karyotype on CVS. An overview of all the samples including gestational age, maternal age and year of sampling can be viewed in the supporting information, S1 Table.

The project was approved by The Regional Committee on Health Research Ethics and The Regional Scientific Ethical Committees for Southern Denmark (Project no: S-20120042). The material used was excess DNA from routine investigation, stored at the biobank at the Clinical Genetic Department at Vejle Hospital. The samples were anonymized and de-identified prior to analysis. The institutional board at the Department of Clinical Genetics and The Regional Scientific Ethical Committees for Southern Denmark therefore waived the need for written informed consent for this study.

Processing of samples was done under identical conditions, however not all at the same time. The samples were not blinded.

DNA extraction and quantification

**Blood samples.** DNA from blood samples were extracted using a standard salt extraction method as described in our previous study [7]

**CVS samples.** DNA from CVS samples was extracted using a QIAamp DNA Mini kit from Qiagen (QIAGEN Inc., Valencia, CA, USA) according to standard protocol provided by Qiagen. The samples were analyzed after separation of maternal decidua and without cell culture.

DNA Quantification

DNA samples were quantified with a Nanodrop ND-1000 (Nanodrop Technologies, Wilmington, DE, USA)

DNA methylation analysis–Infinium microarray analysis

The Illumina Infinium HumanMethylation450 Beadchip Kit (Illumina Inc., San Diego, CA, USA) was utilized for generation of methylation data for all samples. The analysis was done according to standard protocol provided by Illumina. Bisulfite conversion was done using a Zymo Research EZ DNA methylation kit (Zymo Research, Irvine, CA, USA). Beadchips were scanned with an Illumina HiScanSQ scanner using standard settings. Initial quality control, background subtraction and raw data normalization were done using the standard algorithms provided in Illumina Genome studio Methylation module v1.0. Methylation levels are quantified using β-values, as recommended by Illumina. Briefly, the β-value for each interrogated CpG site represents the fraction of methylated versus non-methylated probes, and consequently, the β-value ranges from 0 (unmethylated) to 1 (fully methylated). All analysis and statistical testing was performed on β-values.

Data quality

The Infinium arrays include several control probes for determining data quality, including bisulphite conversion controls. Diagnostic plots of all control probes were visually inspected in the Genome studio software for the approval of each of our arrays. We have added the quality control plots in supporting information(S2 Fig).

Furthermore, 4 samples were analyzed in duplicates on different bead chip arrays. These replicates were analyzed to ensure reproducible data. Methylation data from the different
beadchip showed very strong correlation between replicates \( r > 0.99 \). Validation data for the reproducibility of the 4 replicates can be viewed in S3–S5 Figs and S3 Table.

In addition, the microarray beadchip encompasses 65 probes for highly polymorphic single nucleotide polymorphisms (SNPs). When comparing the 65 polymorphic probes (correspond to genotype) between replicates, the status for each of these 65 sites was highly comparable between replicates (Spearman rho for each pair of replicates: 0.96 to 0.98). \( \beta \)-value of 0.0, 0.5, and 1.0 showed clear distinct patterns of homozygous or heterozygous methylation (Scatterplot of the correlation can be viewed in S6 Fig).

### Bioinformatics

Prefiltering and analysis of identified differentially methylated CpGs (DMCs) was performed as previously described [7], for all methylation profiles. In short, differences in methylation status between sample groups (e.g. CNOR vs T13) were evaluated for each CpG-site using the non-parametric Wilcoxon signed-rank test. The P-values were subsequently adjusted for multiple hypotheses testing using Benjamini–Hochberg correction. The resulting false discovery rates (FDRs) were used in combination with a \( \Delta \beta \)-value cut-off (i.e. a minimum required difference in methylation level) between sample group means to define differentially methylated CpGs (DMCs). The addition of a \( \Delta \beta \)-value cutoff in our definition of DMCs ensures that the observed methylation differences between sample groups (e.g. between the CNOR samples and the T13 samples), for a given CpG, is of a sufficient magnitude to be considered of biological relevance. The full dataset has been deposited in the gene Expression Omnibus (GEO) database (accession number GSE66210). In addition, the GOrilla web tool[10,11] was used for Gene Ontology (GO) analyses. The target and background gene lists were obtained by assigning each CpG to the nearest gene (RefSeq). Consequently, the GO analyses uses all genes represented (i.e. nearest gene) on the 450K array as background, and a small subset based on identified DMCs, as target list.

### Results

#### General methylation landscape

First the general methylation landscape in the different groups was described. Maternal white blood cells (MBCs) showed a very sharp and clear bimodal methylation distribution pattern (Fig 1). Thus, 36% of the CpGs for MBCs were hypomethylated (\( \beta \)-value <0.2) and 39% were hypermethylated (\( \beta \)-value >0.8, Table 1). The CVS samples also showed a somewhat bimodal distribution for the methylation of the CpGs, but the methylation pattern was partially shifted from hypermethylation to semimethylation. Hence, 35% of the CpG sites for normal diploid (CVSNORM) were hypomethylated whereas only 16% were hypermethylated. The general methylation pattern observed for each of the trisomic CVS (T21, T18 and T13) closely resembled that of the normal diploid CVS (CVSNORM), although minor changes in the fraction of hypermethylated sites was observed compared to CVSNORM (Table 1). Thus the fraction of hypermethylated sites for CVS\(_{T21}\) was slightly increased (17.81%), whereas it was decreased for CVS\(_{T13}\) and CVS\(_{T18}\) (13.96% and 12.64%, respectively). No major differences was observed in the methylation pattern for individual chromosomes except for chromosome X in MBC samples, which is probably caused by a substantial methylation of one of the X-chromosomes due to the X inactivation in females (S1 Fig).

Next we explored the methylation pattern of sites mapping to CpG-islands, CpG shores (2 kb flanking CpG islands), CpG-shelves (2kb flanking CpG shores) and open sea (CpGs not mapping to islands, shores or shelves). In general, the methylation pattern of sites mapping to each of these four regions is remarkably different with islands generally hypomethylated,
shores bimodal methylated and shelves and open sea hypermethylated. In addition, placental DNA is generally less hypermethylated (and concomitantly more semimethylated) for all four regions, compared to MBC DNA (Fig 1).

Differences in CpG methylation between DNA from MBCs, normal CVS samples and trisomic CVS samples

Comparing the methylation profiles of MBC and CVSNORM revealed a huge number of DMCs in agreement with the findings shown in Fig 1. When comparing CVSNORM to placental DNA from the different trisomies, only 3 DMCs between CVSNORM and CVST21 was identified, whereas 217 and 219 DMCs was identified for CVST18 and CVST13, respectively (Table 2, FDR < 0.05). However, the mean difference in methylation level (β value) for CVST18 and CVST13 DMCs was in the range 0.2 to ~0.4, suggesting that all of the identified DMCs represent smaller methylation changes between placental DNA from normal and trisomic fetuses. Surprisingly, only 6 DMC’s overlapped between CVST18 and CVST13.

Assessing the distribution of the DMCs in CVST18, we found an enrichment of the DMCs in CpG islands at the expense of shelves and open sea, when compared to the distribution of all CpG sites (Fig 2). DMCs in CVST13 were not enriched in any location. We also investigated

Table 1. CpG methylation distribution.

<table>
<thead>
<tr>
<th></th>
<th>Hypomethylated β &lt; 0.2</th>
<th>Hypermethylated β &gt; 0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of sites</td>
<td>Percent</td>
</tr>
<tr>
<td>MBC</td>
<td>169392</td>
<td>35.89%</td>
</tr>
<tr>
<td>CVSNORM</td>
<td>164847</td>
<td>34.93%</td>
</tr>
<tr>
<td>CVST21</td>
<td>156705</td>
<td>33.63%</td>
</tr>
<tr>
<td>CVST13</td>
<td>163646</td>
<td>34.67%</td>
</tr>
<tr>
<td>CVST18</td>
<td>164301</td>
<td>34.81%</td>
</tr>
</tbody>
</table>

The table lists the frequency of hypermethylated (mean β > 0.8) and hypomethylation (mean β < 0.2) CpG sites within each sample group. Maternal blood cells (MBC), chorionic villus samples (CVS) from normal (CNOR), CVST21, CVST18 and CVST13.

doi:10.1371/journal.pone.0160319.t001
the distance of the CVS\textsubscript{T18} DMCs to the transcription start site (TSS) for nearest gene, to see if the enrichment in CpG island was related to promoter regions, however we did not find any enrichment related to TSS.

Gene ontology of T18 and T13 DMCs

To get an impression of the biological functions of genes closest to the T13 and T18 DMCs, Gene Ontology (GO) software was used. By associating each DMC to the nearest gene, the 219 CVS\textsubscript{T13} and 217 CVS\textsubscript{T18} DMCs were converted to gene-sets used for the GO-analyses. Only 3 DMCs was identified for T21, an inadequate number to establish a gene-set for a GO enrichment analysis. The CVS\textsubscript{T18} gene-set was highly enriched for several GO terms related to “DNA binding” and “transcription factor binding” coupled to the RNA polymerase II transcription (Table 3a). Table 3b lists the 10 most significant enrichments of molecular functions and biological processes for the CVS\textsubscript{T18} DMCs. In contrast, not a single GO term was enriched for the

<table>
<thead>
<tr>
<th>FDR</th>
<th>MBC v CVS\textsubscript{NORM}</th>
<th>T21 v CVS\textsubscript{NORM}</th>
<th>T13 CVS\textsubscript{NORM}</th>
<th>T18 CVS\textsubscript{NORM}</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.05</td>
<td>143951</td>
<td>3</td>
<td>219</td>
<td>217</td>
</tr>
<tr>
<td>&lt;0.01</td>
<td>143717</td>
<td>0</td>
<td>0</td>
<td>113</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td>142788</td>
<td>0</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>&lt;0.0001</td>
<td>139494</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The table lists the number of DMCs that hold a delta $\beta > 0.2$ at different False Discovery Rate FDR. First column compares Maternal blood cells (MBC) to normal CVS samples (CNOR). The following columns compare T21, 18, and 13 CVS samples to normal CVS, respectively.

Fig 2. Stacked bar plot representation of the distribution of CpG sites. First column shows the distribution of all the 471,956 investigated CpG sites on the microarray split into CpG Islands (black), Shores (dark grey) Shelves (light grey) and open sea (White). Second and third columns show the distribution of sites that were significantly higher methylated in T13 and T18 compared to normal CVS samples.
Discussion

Using methylation array technology we have compared the methylation landscape between MBC DNA and placental DNA and furthermore between normal placental DNA and placental DNA from the three common trisomies, with the aim of exploring if these epigenetic differences could improve cffDNA-based non-invasive prenatal diagnosis (NIPD).

A clear difference in the overall methylation pattern between MBC DNA and all four types of placental DNA was observed. Compared to DNA from maternal blood, CVS samples are substantially less hyper-methylated (with a concomitant increased fraction of semi-methylated sites), corroborating recent observations from two studies [9,12].

With regards to the trisomic samples, we observed that the fraction of fully methylated sites were slightly increased in CVST21 and decreased in CVST13 and especially in CVST18 compared to CVSNORM. Exploring this closer by looking at number of sites that were significantly

<table>
<thead>
<tr>
<th>Table 3 A</th>
<th>Molecular functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO Term</td>
<td>Description</td>
</tr>
<tr>
<td>GO:0043565</td>
<td>sequence-specific DNA binding</td>
</tr>
<tr>
<td>GO:0003677</td>
<td>DNA binding</td>
</tr>
<tr>
<td>GO:0000981</td>
<td>sequence-specific DNA binding transcription factor activity</td>
</tr>
<tr>
<td>GO:0003700</td>
<td>sequence-specific DNA binding transcription factor activity</td>
</tr>
<tr>
<td>GO:0001071</td>
<td>nucleic acid binding transcription factor activity</td>
</tr>
<tr>
<td>GO:0003676</td>
<td>nucleic acid binding</td>
</tr>
<tr>
<td>GO:0044212</td>
<td>transcription regulatory region DNA binding</td>
</tr>
<tr>
<td>GO:0000975</td>
<td>regulatory region DNA binding</td>
</tr>
<tr>
<td>GO:0001067</td>
<td>regulatory region nucleic acid binding</td>
</tr>
<tr>
<td>GO:0001228</td>
<td>RNA polymerase II transcription regulatory region sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3 B</th>
<th>Biological processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO Term</td>
<td>Description</td>
</tr>
<tr>
<td>GO:0045944</td>
<td>positive regulation of transcription from RNA polymerase II promoter</td>
</tr>
<tr>
<td>GO:1903508</td>
<td>positive regulation of nucleic acid-templated transcription</td>
</tr>
<tr>
<td>GO:0019219</td>
<td>regulation of nucleobase-containing compound metabolic process</td>
</tr>
<tr>
<td>GO:0048518</td>
<td>positive regulation of biological process</td>
</tr>
<tr>
<td>GO:0045893</td>
<td>positive regulation of transcription, DNA-templated</td>
</tr>
<tr>
<td>GO:0045935</td>
<td>positive regulation of nucleobase-containing compound metabolic process</td>
</tr>
<tr>
<td>GO:0006357</td>
<td>regulation of transcription from RNA polymerase II promoter</td>
</tr>
<tr>
<td>GO:0006355</td>
<td>regulation of transcription, DNA-templated</td>
</tr>
<tr>
<td>GO:0051252</td>
<td>regulation of RNA metabolic process</td>
</tr>
<tr>
<td>GO:1903506</td>
<td>regulation of nucleic acid-templated transcription</td>
</tr>
<tr>
<td>GO:0051173</td>
<td>positive regulation of nitrogen compound metabolic process</td>
</tr>
</tbody>
</table>

A: Gene functions related to terms within the group “Molecular functions”. B. Gene functions related to terms within the group “Biological processes”. The table shows the enrichment analysis for the genes associated to our DMCs in T18 samples. The total number of genes within the GO analysis is 17727. Our gene set related to the CVS T18 DMCs were 118. Each table shows the 10 GO terms with the lowest False Discovery Rate (FDR q value).

doi:10.1371/journal.pone.0160319.t003

CVS_{T13} gene-set. The full list of enriched biological processes can be viewed in the supplementary material (S2 table).

Microarray-Based Analysis of Methylation of 1st Trimester Trisomic Placentas

PLOS ONE | DOI:10.1371/journal.pone.0160319 August 4, 2016 7/11
different methylated in CVS \textsubscript{T21}, CVS \textsubscript{T13} and CVS \textsubscript{T18} DNA compared to CVS\textsubscript{NORM} DNA we observed very few DMC’s in CVS \textsubscript{T21} DNA and in contrast a substantial and very similar number of DMC’s for CVS \textsubscript{T13} and CVS \textsubscript{T18} DNA but with a very limited amount of overlap. The DMC’s in T18 were more robust—lowering the FDR (increasing level of significance) lead to clear differences between T18 and T13 (see Table 2). However, it should be remembered that the sample sizes were different since, we could only include six T13 samples due to the rarity of T13 pregnancies. We do not know the explicit reason as to why T13 and T18 have a higher number of DMCs compared to T21, however it does coincide with the more severe phenotypes for T13 and T18 in which the affected foetuses often die in the uterus or within the first few years of life as compared to T21 affected foetuses.

Subsequently Gene Ontology enrichment analysis showed for CVS \textsubscript{T13} DMC’s that among the gene functions of the nearest genes, there was no significant overrepresentation of specific functions compared to the reference gene set. In contrast, among the functions of the nearest gene to the CVS \textsubscript{T18} DMC’s, there was a highly significant overrepresentation of gene functions related to regulation of RNA polymerase II mediated transcription and sequence-specific DNA binding related to transcription factor activity. The smaller number of T13 samples does, however, increase the risk of false negative findings in the GO analysis, but in spite of this we find the difference in significant GO terms between CVS \textsubscript{T18} and CVS \textsubscript{T13} interesting. We have at present no explanation for this difference but as DNA methylation to some extent can be considered a weak proxy for gene expression it would be interesting to investigate if there are differences in gene expressions of genes involved in regulation of RNA polymerase II mediated transcription between CVS \textsubscript{T18} and CVS \textsubscript{T13} samples.

We are aware that caution should be taken in the interpretation of microarray based data from CVS samples, because the chorionic villi is a heterogeneous mixture of syncytiotrophoblastic-, cytotrophoblastic-, mesodermal- and fetal endothelial/vascular cells, and the proportion of the different cell populations in the biopsy could be a confounding factor. This is, however, to our knowledge the first study to date to investigate the methylation landscape in T13 and T18 placentas. The relatively large number of samples, especially for T13 and T18 placentas, minimizes the risk of random variability and therefore provides a more representative biological measure. To further limit the risk of confounding factors such as sex and gestational age we have used gestational–age-matched blood samples with an equal distribution of samples with male and female fetuses.

Regarding the circulating cell free DNA (cfDNA) we choose to use maternal white blood cells as a proxy, since cfDNA in plasma from non-pregnant individuals predominantly originates from hematopoietic cells [13]. Therefore, it is assumed that the maternally derived cfDNA in blood from pregnant women has the same origin. However, one study has suggested that for very obese women a substantial fraction of adipose apoptotic DNA could be released into the maternal systemic circulation [14]. Optimally maternal cfDNA should have been used. However, the very low cfDNA fraction in blood samples hinders its appliance for microarray analysis.

Only very few publications have until now looked at methylation in trisomic placental DNA. However, the observation of a small but significant increase in fraction of hypermethylated CpG sites in placental samples from T21 cases compared to samples of normal placentas, are in line with two recent publications showing a general hypermethylation across all chromosomes in fetal DNA from T21 placentas [8,9]. Furthermore Jin et al. found several of the epigenetic changes to be conserved across tissues and they proposed that a down regulation of a group of the \textit{TET}-family genes known to play an important role in epigenetic regulation could be responsible for the overall hypermethylation seen in all chromosomes through decreased DNA demethylation[9,15]. However, we identified only three CVS \textsubscript{T21} DMCs using the most
lenient criteria (delta \( \beta > 0.2 \) & FDR < 0.05), suggesting that the methylation differences are either relatively small or not very consistent. Eckmann-scholz et al. found 464 DMC between normal and T21 CVS samples, with a delta \( \beta > 0.2 \). They applied the same Illumina methylation arrays as in our study, however only assessing 27,000 CpGs and the number of T21 samples in their study was only 3[8].

Assessing the overall methylation landscape in relation to the distance to CpG islands, we found the most highly different methylated CpGs in T18 to be enriched in CpG islands and decreased in open sea and shores, when compared to the distribution of all the covered CpGs. However, when investigating the distribution of CVS\textsubscript{T18} DMC in relation to transcription start sites (TSS), we did not find any enrichment of the DMCs in TSS vicinity, indicating that the CVS\textsubscript{T18} DMCs are located at non-promotor associated CpG islands.

We have in the present communication provide a detailed methylation analysis of samples from T13, T18 and T21 placentas and compare them to normal placenta tissues and maternal blood. We found a substantial number of significant CpG methylation differences in T13 and T18 placenta DNA compared to T21 placenta DNA. Our data suggests that the genes associated with CVS T18 DMCs are enriched for biological pathways/processes related to RNA polymerase II mediated gene expression. Our findings do not support the idea of using differences in the specific methylation imprint (or landscape) of T13, T18 or T21 as targets for NIPD.

**Supporting Information**

**S1 Fig.** Density plot of methylation level distribution for CpG-sites on different chromosomes. X axis represents methylation level as mean \( \beta \)-values. Y axis represents relative density. Maternal blood cells (MBC) red line, CVS normal blue line, CVS T13 black line, CVS T18 dark grey line, T121 light grey line. (EPS)

**S2 Fig.** Diagnostic quality control plots. The Illumina Infinium HumanMethylation450 Beadchip arrays include several control probes for determining data quality, including sample-independent controls: staining controls, extension controls, target removal controls and hybridization controls. Sample-dependent controls: Bisulphite conversion I controls, Bisulphite conversion II controls, Specificity I controls, and Specificity II controls, Nonpolymorphic controls and Negative controls. Diagnostic plots of all control probes, visualized by illumines genome studios software, are presented for each of the 3 beadchip arrays. (ZIP)

**S3 Fig.** Validation of DNA methylation data reproducibility. Four replicate samples, two maternal blood samples (MBC) and two CVS-samples, underwent independent bisulphite conversion and were analyzed on different beadchip arrays for the validation of DNA methylation data reproducibility. The figure shows hierarchical clustering and heatmap of correlation coefficients based on all 472K methylation sites. The four replicates MBC3, MBC6, CNOR3 and CNOR5 all shows the highest correlation between all samples. (PDF)

**S4 Fig.** Density scatter plots of the four replicates. All four replicates have a correlation coefficient > 0.99 for all methylation sites. (JPG)

**S5 Fig.** Reproducibility and completeness of bisulphite conversion of unmethylated sites. The figure shows the mean delta \( \beta \) values for all unmethylated sites (average \( \beta < 0.2 \)) for the four replicates. Since the replicates underwent independent bisulphite conversion, complete
bisulphite conversion of unmethylated sites should result in reproducible low beta values. A. Delta β values for all unmethylated sites for the two maternal blood cell (MBC) replicate samples. Sorted with decreasing delta β values. X axis represents the numbers of unmethylated sites. B. The distribution of the 100 highest delta β values for the MBC samples. C. Delta β values for all unmethylated sites for the two CVS replicate samples (CNOR). D. The distribution of the 100 highest delta β values for the CVS samples.

S6 Fig. Scatterplot displaying data reproducibility of highly polymorphic single nucleotide polymorphisms (SNPs). The beadchip encompass 65 highly polymorphic single nucleotide polymorphisms (SNPs) The scatterplot displays the β-values for 65 polymorphic SNPs for the four replicate samples. Spearman correlation coefficient (p) for each pair of replicates is shown. Diagonal red line represents complete correlation (p = 1.00). Methylation levels (β-value) of 0.0, 0.5, and 1.0 shows clear distinct patterns of homozygous or heterozygous methylation.

S1 Table. Overview of all samples used in the methylation analysis. The table lists samples, fetal gender, year of sampling, sample material, chromosome analysis, gestational age (weeks), and maternal age (in years).

S2 Table. Gene Ontology (GO) analyses related to CVS T18 DMCs. Gene functions related to terms within the group “Biological processes”. The table shows the enrichment analysis for the genes associated to our DMCs in T18 samples. The total number of genes within the GO analysis is 17727. Our gene set related to the CVS T18 DMCs were 118.

S3 Table. Overview of the methylation differences between replicates. Numbers (and percentages) of sites with increasing delta β value thresholds. A All methylation sites. B Unmethylated Sites (average delta β < 0.2).

Acknowledgments
Corresponding author, Lotte Hatt, accepts responsibility for the integrity and validity of the data collected and analyzed.

Author Contributions
Conceptualization: LH SK AB CB IEA SS JG MMA.
Data curation: AB JG MMA.
Formal analysis: JG MMA.
Funding acquisition: LH SS SK AB.
Investigation: LH JG MMA.
Methodology: AB JG MMA.
Project administration: LH AB SK.
Resources: IEA CB SS.
Software: JG MMA.
Supervision: SK AB CB IEA SS.
Validation: LH JG MMA.
Visualization: LH SK AB CB IEA SS JG MMA.
Writing - original draft: LH.
Writing - review & editing: SK AB CB IEA SS JG MMA.

References