Epigenetic Profiling of Human Placenta Throughout Early Gestation

Qianhui Wan

Supervisors: Prof. Claire Roberts,

Dr. Tina Bianco-Miotto and Dr. James Breen

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Medicine

Adelaide Medical School

Discipline of Obstetrics and Gynaecology

University of Adelaide

March 2020

Declaration of Authorship

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed: ___ ___

Date: 27th of March 2020

Abstract

Epigenetic Profiling of Human Placenta Throughout Early Gestation

Qianhui Wan

The differentiation of the placenta, especially during early gestation is important for pregnancy success. Whilst emerging evidence has shown that DNA methylation (DNAm) in placenta varies over gestation, to date, most studies have compared DNAm in a relatively short gestational age range. Little is known about the dynamics of DNAm patterns across early pregnancy from as early as 6 weeks', and as late as 23 weeks' gestation. This comprehensive analysis of the placental methylome will help to elucidate the previously poorly understood relationship between DNA methylation, placental development, and complications of pregnancy. The overall aim of this thesis is to characterise and interpret this relationship through the analysis of DNA methylation profiles from strictly phenotyped samples of human placenta, and matched maternal leukocytes, across early to mid-gestation, through bioinformatics analyses. All data herein were obtained using the Illumina Infinium® MethylationEPIC BeadChips (EPIC arrays).

In this study, we first compared three different bioinformatics methods implemented with different algorithms for detecting differentially methylated regions (DMRs) between sample groups. Subsequent to these analyses we aimed to establish an in-house pipeline for the quality control and analysis of EPIC array methylation data obtained from both GEO database, and from this study. The three methods used for the discovery of DMRs were bumphunter, Probe Lasso and DMRcate. After comparison of these three methods we were able to demonstrate unique advantages

and disadvantages of each. Overall, DMRcate was considered the most appropriate method for the identification of DMRs in EPIC array methylation data from our placenta samples, with a better sensitivity than Probe Lasso and bumphunter methods and less false positive regions than the Probe Lasso method.

Next, the established in-house pipeline was used for array data analyses. Initial unsupervised clustering using a PCA analysis of methylation data revealed several outliers within our data. These 6 samples did not cluster as expected with other placenta samples of the same gestational age. To investigate whether these outliers were caused by complicated pregnancy or technical issues, we compared the data from samples in our study with publicly available data from samples of placenta and placenta-associated tissues. Given the otherwise strong gestational age clustering observed in the PCA analysis, and the unknown pregnancy outcomes of the tissue in question, we hypothesized that the samples which failed to cluster within their gestational age group would cluster with other like samples. After preprocessing we included the public data in a new PCA analysis with results indicating that the outliers we identified were not pure placenta villous tissue, but rather these samples were a mix of both placental and maternal tissue.

After assessing the quality of all placenta samples, and removing samples identified as containing maternal tissue, an epigenome wide DNAm study of placenta (n = 125) across 6-23 weeks' gestation was performed. Placental DNA methylation changed throughout gestation, with methylation differences also found between groups up to and after 10 weeks' gestation. Since maternal blood starts to flow into placenta at approximately 10 weeks' gestation, these DNA methylation changes could be

associated with a change in oxygen tension in the placenta. Further to the DNA methylation changes identified across early gestation, DNA methylation levels at partially methylated domains and imprinting control regions were stable in placenta across early gestation, suggesting an association with these regions and the basic function and development of the human placenta.

Finally, DNA methylation changes of maternal leukocytes from matched maternal blood were investigated. We identified DNA methylation changes in maternal leukocytes associated with maternal smoking and with maternal age, and to a lesser degree we were able to identify changes in DNA methylation of maternal leukocytes that were associated with gestational age. Changes of cell proportion for maternal leukocytes were identified and a potential accelerated aging was found in pregnant women compared with non-pregnant women. These findings provide more information for real time assessment of pregnancy health using DNA methylation in maternal circulating leukocytes.

In summary, the research reported here provides an insight into performing bioinformatics analyses and quality control of placental DNA methylation data obtained from EPIC array analyses. Further, this thesis adds to our understanding of placental development, health and disease through the characterisation of the DNA methylome of placenta and matched maternal leukocytes across early gestation.

Acknowledgements

This work would not have been possible without the continuous support from my supervisors. I would like to first sincerely thank my supervisors Claire Roberts, Tina Bianco-Miotto and James Breen, for spending time guiding me with great patience throughout my PhD and making my PhD journey a wonderful and rewarding experience in my life. It has been a great honor and very lucky to work with scientists who have a passion and a love of medical and clinical research. In particular I would like to express my thanks to my primary supervisor Claire Roberts for having me as a PhD student in The Placenta Lab, and for her generous provision of guidance on placental biology and for her care and encouragement throughout. I am very grateful for the guidance provided on epigenetics from Tina Bianco-Miotto, and for the knowledge and skills of bioinformatics from James Breen.

I would like to acknowledge all the members of The Placenta Lab including Shalem Leemaqz who helped me with statistical problems, Dylan McCullough, Dale McAninch and Tanja Jankovic-Karasoulos who collected all the samples and undertook the laboratory work, and lab members including postdoctoral fellows Prabha Andraweera, Luke Grzeskowiak, Jessica Grieger and visiting post-doctoral researcher Callista Mulder who spent time discussing questions associated with my project, and post-graduate students Melanie Smith, Justin Bogias, Amy Garrett, Michelle Plummer, Nahal Habibi, Maleesa Pathirana, Benjamin Mayne and Rebecca Wilson who always encouraged me and cheered me up during my PhD. I also want to express my heartfelt thanks to all the members in the bioinformatics hub who assisted me with my coding and scientific writing.

I would like to acknowledge the funding bodies that financially support this project including China Scholarship Council (CSC) for awarding the scholarship for my PhD, the grants received by The Placenta Lab especially the project grant from National Institute of Child Health and Human Development (NICHD) and National Institutes of Health (NIH). I am very grateful to my supervisor Claire Roberts for providing me with a scholarship during the 6 months' extension of my PhD candidature.

I also want to thank my mentor Richard Preston (1963-2017) in the "Talking with Aussies" program which is a volunteer program aimed at helping international students to improve their English conversation skills in the University of Adelaide. Richard spent much of his time doing different kinds of volunteer teaching to educate students, including me, without letting us know that he was diagnosed with late-stage pancreatic cancer in 2015 and suffered great pain because of this disease. I must thank him for the selfless gift he gave both as a great educator and as a role model and sincerely thank him for helping me to understand the true culture of Australia.

I would like to thank the colleagues and staff from the University of Adelaide for consistently helping me throughout my PhD including all the members from Robinson Research Institute, Dr Paul Bolton from University Health Practice, Joanne Mcnamara from Student Life Counselling Support, Jo Porter from Student Care Office, staff from the Adelaide Graduate Centre and International Student Support.

I would like to thank the friends I met in Adelaide including Yutong Liu, Rong Hu, Ning Liu, Jing Zhao, Yang Zhao, Yanchen Zhang, Bihong Zhang, Doan Dinh, Monica Hartanti, Nicole Bastian, Jason Gummow, Siew Siew Lee, Linda Wu, Siew Leng Wong,

Liang Chen, Yifei Huo, Xuening Cui, Bo Li and my friends that I have met in China including Xiaoli Yu, Fan Yang, Yi Cui, and Jinglong Zhang for supporting me and helping me getting through the problems in my work and life.

I have also received a lot of encouragement from New Life Christian Community (NLCC) Church, I must thank all the members in this church especially Rachel Lo, Qing Fang, Fiona Zhao, Hanci Amelia Jiang and Catherine Hao Yue Poon for supporting me and sharing their wisdom with me.

Different kinds of music are full of creativity and are my important companion during my PhD, so I need to thank musicians including Wolfgang Mozart, Ludwig Beethoven and also the rock band BEYOND and the rock music composers Yida Huang and Crowd Lu.

At last I want to mention that I am very grateful for the support of my family members including my father, mother, younger sister, cousins, grandparents and other relatives. Their unconditional love is priceless and will be my motivation forever.

Publications Arising from this Thesis

1. **Qianhui Wan**, Shalem Yiner-Lee Leemaqz, Stephen Martin Pederson, Dylan McCullough, Dale Christopher McAninch, Tanja Jankovic-Karasoulos, Melanie Denise Smith, Konstantinos Justinian Bogias, Ning Liu, James Breen, Claire Trelford Roberts and Tina Bianco-Miotto. Quality control measures for placental sample purity in DNA methylation array analyses. ***Placenta*** 2019, 88:8-11.

2. **Qianhui Wan**, James Breen, Shalem Yiner-Lee Leemaqz, Stephen Martin Pederson, Tanja Jankovic-Karasoulos, Dale Christopher McAninch, Dylan James McCullough, Melanie Denise Smith, Konstantinos Justinian Bogias, Ning Liu, Callista Leonie Mulder, Claire Trelford Roberts and Tina Bianco-Miotto. DNA methylation profiling of human placenta samples across early gestation. ***To be submitted June 2020***

Contents

List of Figures

List of Tables

1 Introduction

The placenta plays a vital role in fetal development and maternal health since it influences not only the health of a woman and her fetus during pregnancy, but also the lifelong health of both mother and child [1]. Hormones and cytokines released from the placenta have key effects on both maternal and fetal physiology and the development of the fetus depends on the maternal circulation for exchanging nutrients and waste products through the vasculosyncytial membrane of placenta [2]. Due to close association with the fetus, the health status of the placenta can also predict the long-term health status of the child, such as the risk of chronic diseases including hypertension and heart diseases in their later life [3].

The placenta can also protect the fetus from adverse factors in maternal blood. The vasculosyncytial membrane is a selective barrier including the syncytiotrophoblast and the endothelial cells of the capillaries, which can prevent some macromolecules from reaching the fetus [4]. As a natural state of maternal insulin resistance, pregnancy induces a high glucose concentration gradient from maternal to fetal blood which facilitates placental, hence fetal, glucose uptake across gestation [5]. This could be associated with the increased number of insulin receptors on the placental villous membrane during early gestation compared to late gestation [6]. In addition, immune tolerance in pregnancy is important to prevent maternal immunological rejection of the fetus and placenta. Disrupted immune function may be associated with placental oxidative stress or excessive inflammation in pregnancy complications [7].

The DNA methylation profile of human placenta is unique as it is globally hypomethylated compared to other non-cancerous, healthy tissues [8]. Apart from global hypomethylation, placenta is hypermethylated at CpG dense regions and placental DNA also contains large (megabase-scale) low methylated regions, known as partially methylated domains (PMDs) that cover about 37% of the placental genome [9]. These PMDs do not exist in other normal tissues but are commonplace in epithelial cancers such as colon and breast cancer [10]. Studies have shown that placental DNA methylation increases across gestation and anomalous variations are associated with poor pregnancy outcomes [11]. For instance, abnormally methylated promoters of genes such as *LRAT*, *SLC19A1*, *EFS* and *SR140* are associated with intrauterine growth restriction (IUGR) which is a condition related with poor fetal growth during gestation [12]. Fully characterising the dynamic and changing DNA methylation of human placenta across gestation is important for our understanding of placental development, and of how DNA methylation affects pregnancy outcomes.

DNA methylation can dramatically impact the way that genes are expressed, most notably through the accumulation of methylated cytosines at the promoters of genes. Currently the most efficient way to identify these DNA methylation sites, as well as their effect on gene expression, is through the application of genomic approaches such as high-throughput sequencing and array technologies. RNA sequencing (RNA-seq) can identify novel variations in gene expression, non-coding RNA species and splicing variants, which may be important for understanding the regulatory mechanisms of normal or pathological tissue development [13-15]. Combined with RNA sequencing data, the analysis of DNA methylation array data can reveal the sites of DNA

methylation that may be differentially modified in different conditions and associated with gene expression.

Sequencing, both for RNA and DNA, and microarrays generate a large amount of data [16]. Analysis of this massive amount of data through bioinformatics methods is essential to interpret the association between biological phenotypes and molecular changes. It may provide very useful clues to guide future experiments. In this chapter, we review the mechanisms and functions of DNA methylation in placenta and maternal blood during human pregnancy, highlight the importance of characterising DNA methylation profiles during pregnancy and introduce and compare methods for analysing DNA methylation data.

1.1 Development and function of human placenta

The placenta is essential for fetal development in eutherian mammals and abnormal placental development is associated with pregnancy complications such as preeclampsia (PE), a maternal hypertensive disorder, and IUGR [17]. There is great variation of the types of placenta between different mammals, which indicates that the function of animal models in understanding the development of human placenta is limited. Here, we reviewed the development and function of human placenta with a main focus on the placenta chorionic villus which is the functional unit of human placenta [18].

1.1.1 The development of the placenta

Successful human pregnancy and fetal health are closely related to the development of the placenta, the temporary exchange organ between the mother and her fetus. The placenta is an organ that develops from the fetal trophectoderm (diverse trophoblast cell types) and inner cell mass (blood vessels, mesenchyme, chorion, and amniotic membranes) [19, 20]. The development of the placenta begins from as early as the end of the first week after fertilisation when the blastocyst starts to implant into the uterus [21]. The implantation of the embryo is a kind of erosion process, which begins with the differentiation of trophoblasts at the embryonic pole [22]. The trophoblasts differentiate into syncytiotrophoblast (STB) and cytotrophoblast (CTB). CTBs fuse to form STB, a multinuclear and terminally differentiated cell type [23, 24]. With the development of the blastocyst, the multinuclear STB expands rapidly, continuing its invasion of the endometrium and consequently embeds the conceptus completely into the endometrium, termed decidua in pregnancy, at the end of the second week post conception. Concurrently, the lacunae containing maternal venous fluid, form between folds in the STB and later fuse to form the lacunar network which lays the foundation for the development of the intervillous space [25].

Chorionic villi are very important for placental function because they are the interface between the maternal and fetal circulations where the exchange of gases, nutrients, wastes and hormones occurs [26]. Importantly, abnormal villous development has been associated with pregnancy complications such as preeclampsia and IUGR [27]. Primary chorionic villi first become recognisable between the second week and early third week of gestation as CTB cells penetrate into the cords of the STB. Shortly after

this, mesenchymal cells (extra-embryonic mesoblast) migrate into the primary villi to form mesenchymal tissue cores, now known as secondary chorionic villi, expand into the lacunae that, as we mentioned earlier, are filled with maternal venous fluid [28]. Mesenchymal cells may also differentiate into the blood cells and capillaries. Villi that contain these differentiated blood vessels are known as tertiary villi and are seen prior to week 6 of gestation when the blood vessels are visible. Arteriocapillary networks are formed through the fusion of capillaries in the villi and later these capillaries will be connected to the embryonic heart [29].

In early development, the outer layer of the blastocyst (the trophectoderm) is comprised of trophoblasts that can develop into the chorionic membrane. At approximately eight weeks' gestation, villi opposed to the decidua begin to regress, undergoing almost complete degeneration, and form the chorion laeve. Villi in contact with the decidua basalis however, proliferate and form the fetal component of the placenta [2] known as the chorion frondosum, by the end of the first trimester [18, 19, 30]. When these villi disappear, the villi associated with the decidua basalis (villous chorion) increase rapidly in number [2]. Along with the growth of the fetus, the size and thickness of the placenta increases until approximately 24 weeks' gestation. Placental growth decreases in the last trimester however, it has been shown that approximately 15 percent of placentas show a continuous increase in weight throughout pregnancy [31].

Terminal villi provide a large surface area for the exchange of various substances between the maternal and fetal circulation. These terminal villi are bathed in maternal blood which fills the intervillous space [32]. The fully developed placenta covers 15%

to 30% of the decidua and weighs about one sixth of the weight of the fetus. After a successful pregnancy, the placenta is delivered after delivery of the fetus, between 37-41 weeks after fertilization [2]. The developmental processes of the placenta and the fetus are shown in Figure 1.1.

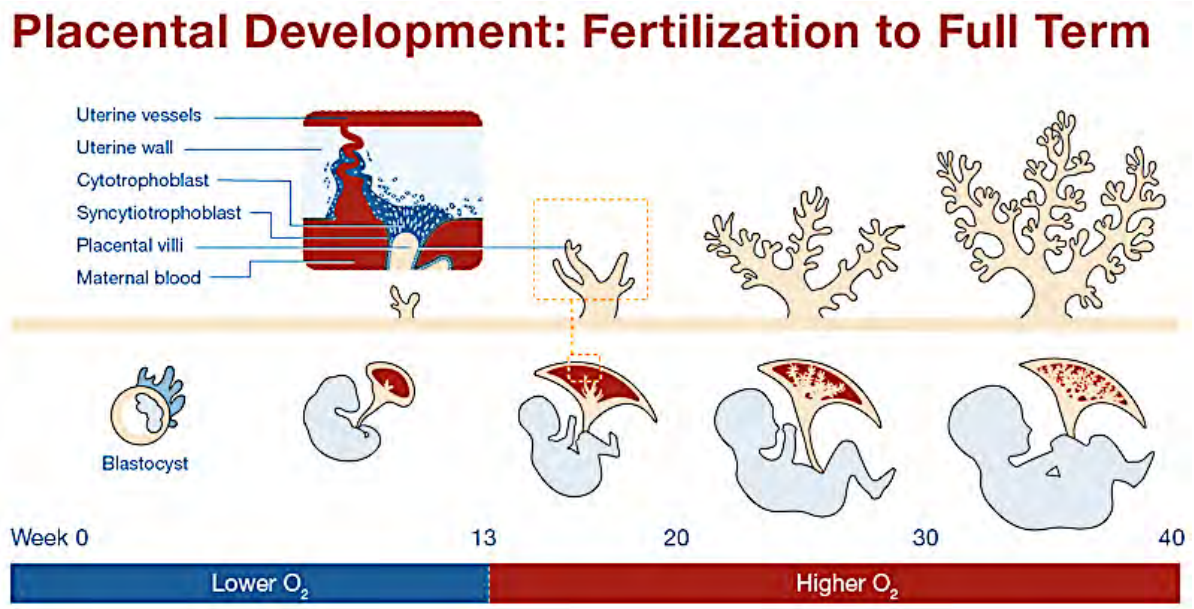

Figure 1.1 The timeline shows the development of the placenta and fetus across 40 weeks of gestation and the corresponding relative oxygen concentrations [33].

During placental development the oxygen environment changes (Figure 1.1). This change in oxygen tension plays an important role in a number of mechanisms that associate with normal placental function. Placental cells have a sensing system for hypoxia, nutrient supply and energy supply through HIF, mTOR and AMPK pathways, respectively, which are critical for maintaining maternal and fetal metabolism during pregnancy [34]. During early gestation, there is no maternal blood flow into the intervillous space, so early placentation occurs in a relatively hypoxic environment which has been shown to be important for early placenta development [35]. The hypoxic environment of less than 20 mmHg until 10-12 weeks' gestation in placental

tissue is due to the cytotrophoblast occlusion of the spiral arterioles [36]. Oxygenated maternal blood starts to flow into the intervillous space of the placenta around 10-12 weeks' gestation, resulting in an increase in oxygen tension (PO_2) (Figure 1.2) [37]. Studies show that this changing PO_2 is related to trophoblast invasion and vascular remodelling. Hypoxia induces the invasion and migration of trophoblasts which is associated with the pathogenesis of PE [38]. If the invasion of trophoblast is reduced, placental blood flow and oxygen availability will also be reduced [39], which may cause a release of soluble factors into the maternal blood which can lead to PE [40].

Though the placenta is of great importance for reproductive success, the understanding of human placenta is still very limited. The great diversity of placenta between species and the ethical obstacles for studying human placenta during early gestation has made the investigation of human placenta development difficult [18, 41].

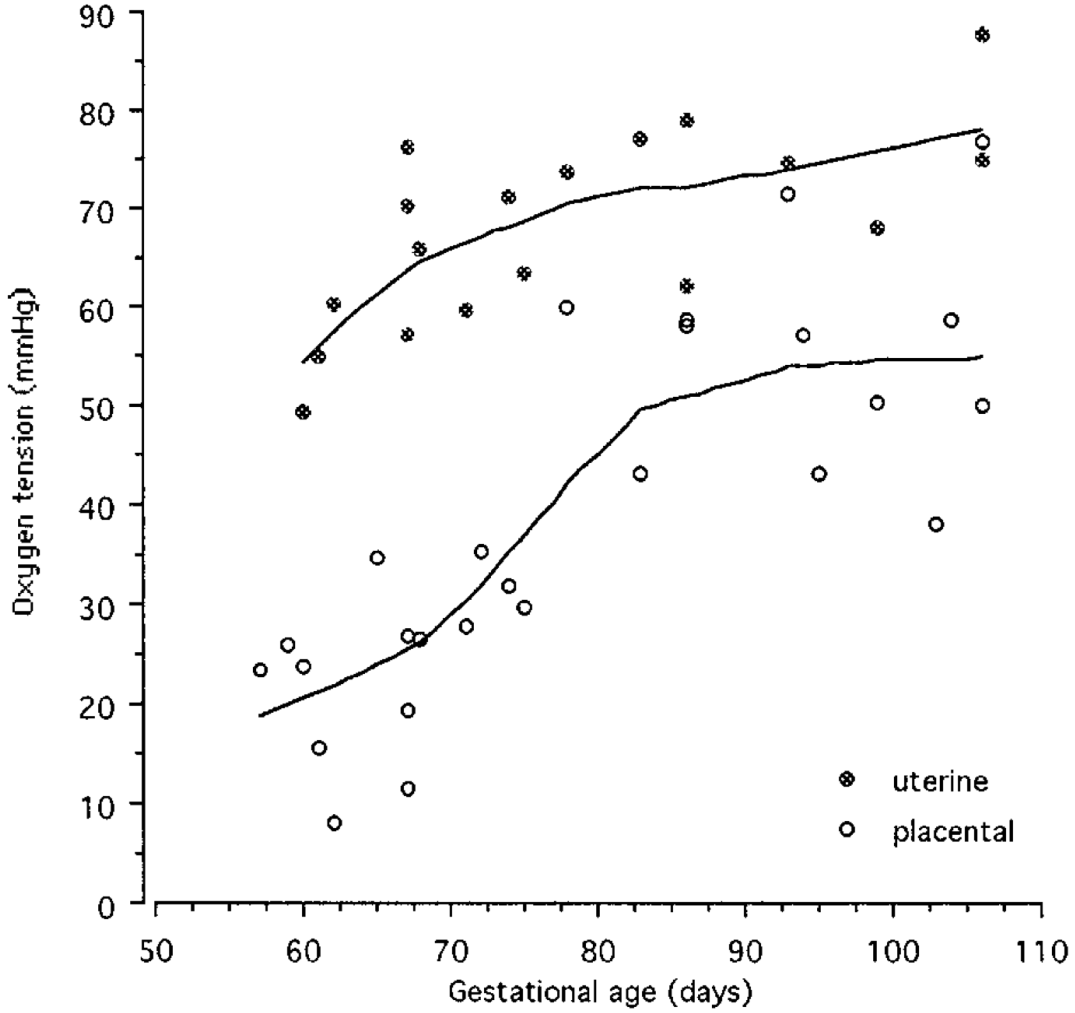

Figure 1.2 PO_2 in placenta and uterine increased across early gestational age [37].

1.1.2 The function of the placenta

During fetal development, the placenta functions not only to provide oxygen and nutrients to the fetus but also in place of important fetal systems such as the fetal respiratory, endocrine and immune system [42]. The placenta anchors the conceptus to maternal tissues and contributes to the immunological tolerance to paternally derived antigens of the growing fetus [43]. Some maternal antibodies (mainly IgG) are

transported through the placenta by transcytosis and consequently passive immunity is obtained by fetus [21, 44]. The placenta produces glycogen, cholesterol and fatty acids that are important for fetal metabolism. Water and gas (e.g. oxygen and carbon dioxide) are transported through the vasculo-syncytial membrane by simple diffusion. Specific transporters are involved in the transportation of glucose, amino acids, and fatty acids. In addition, the placenta produces important proteins and steroid hormones for maintaining pregnancy including human chorionic gonadotrophin (hCG), progesterone, and estrogens [45]. Important placental functions involved in maternal and fetal resource allocation are shown in Figure 1.3.

Disruption of placental development and placental dysfunction are associated with maternal and fetal diseases [46, 47]. Impaired remodelling of the uterine spiral arterioles increases the probability of developing PE which can lead to eclampsia, one of the leading causes of maternal death [48]. A reduction in the circulating maternal blood in placenta has been shown to result in hypoxia and has been associated with IUGR of the fetus [49]. In addition, poor placental development and function during pregnancy is a predictor of disease susceptibility for later life. For example, fetuses with abnormal placental histology tend to have an elevated risk for autism [9]. As the least studied organ compared with other human organs [50], more studies of human placental development at a molecular level are needed.

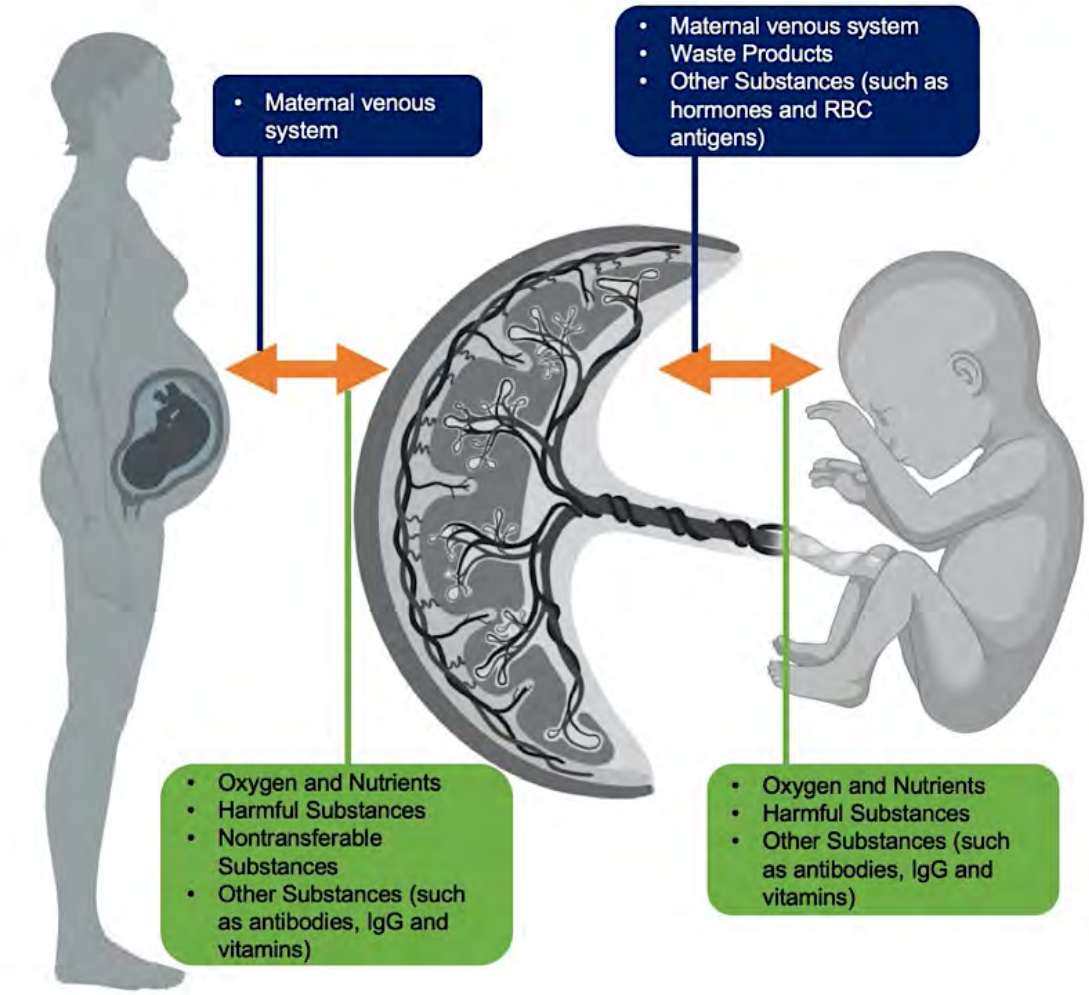

Figure 1.3 The exchanging function of placenta, modified from [45].

1.2 Pregnancy

Many factors such as environmental factors, fetal sex, maternal smoking and BMI are related to the health of pregnant women and their fetuses [51]. Up to ten percent of the pregnancies worldwide are affected by PE which has great impact on maternal and fetal health [52]. Approximately 76,000 maternal and 500,000 infant deaths worldwide were caused by PE each year, which makes PE one of the leading causes of pregnancy related deaths in the world [53]. Currently there are very few effective

treatments for PE except for delivery of the placenta, resulting in babies born too soon and these babies are more likely to have hypertension, cardiovascular disease and diabetes later in their life [54]. Finding biomarkers in maternal blood that can detect preeclampsia or other complications early in pregnancy will be a good way to intervene and monitor the women and the babies early.

1.2.1 Pregnancy complications

Some women can experience pregnancy complications which can affect the health of these women and their fetus. Even though the women are healthy before pregnancy, they may develop complications during their pregnancies [55]. PE, IUGR, small for gestational age (SGA), preterm birth (PTB) and gestational diabetes (GDM) are some commonly diagnosed complications during pregnancies.

Preeclampsia is a serious pregnancy complication characterised by various symptoms including hypertension, proteinuria, kidney injury, liver dysfunction and fetal growth restriction [56]. PE is still one of the main causes of maternal death and morbidity. The pathology of PE has not yet been fully understood, but there is evidence that shows that the pathology of PE is associated with abnormal placentation, remodelling of maternal spiral arteries (Figure 1.4) and maternal immune response [57].

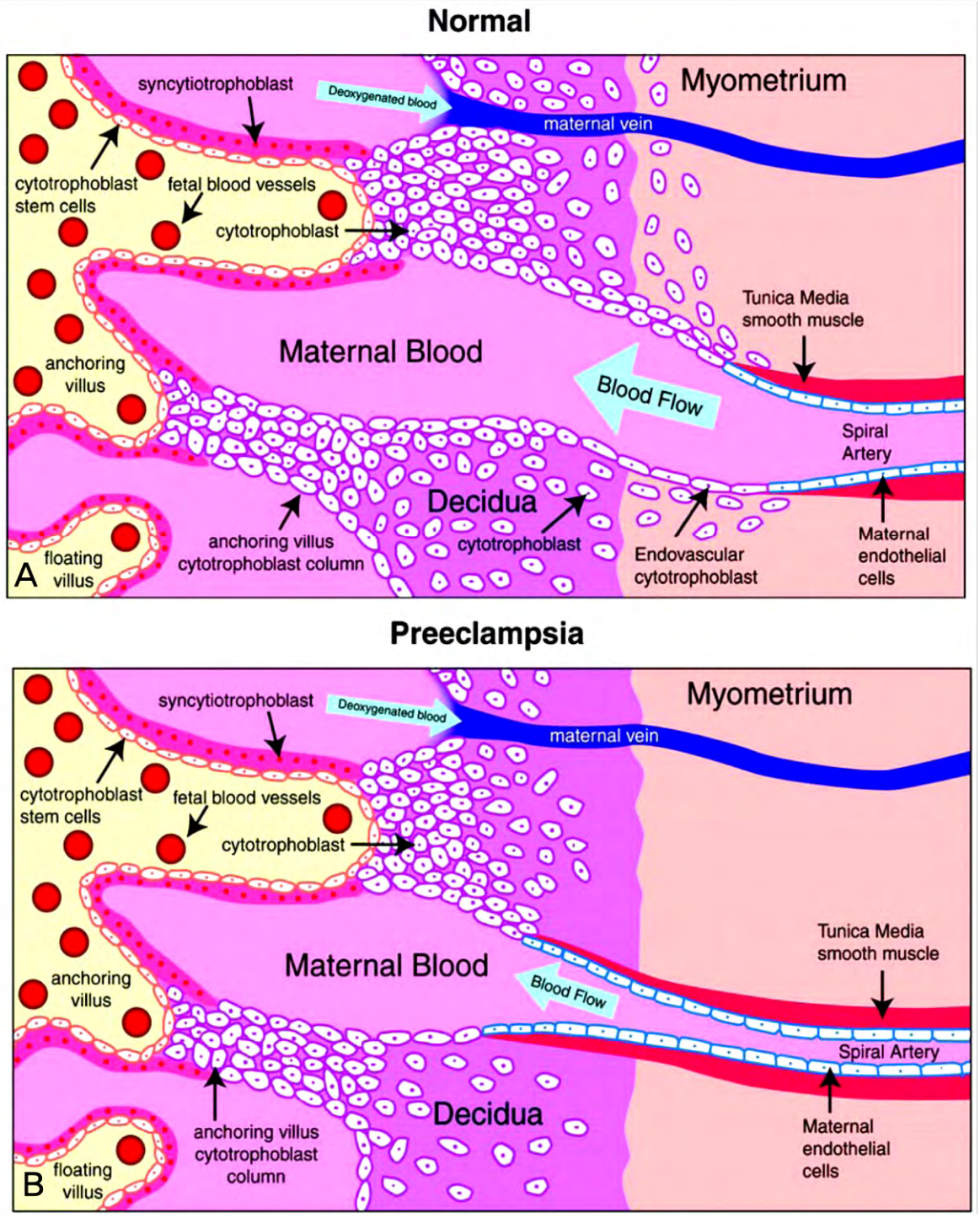

Figure 1.4 The difference of normal placentation and abnormal placentation associated with preeclampsia [58].

Small for gestational age birth is defined as birthweight < 10th centile for gestational age [59] and it can cause fetal and neonatal death. In many cases, the aetiology of

SGA is unknown, however, there are some reported factors that may cause SGA, including pregnancy-induced hypertension, fetal chromosomal abnormality and placental structural abnormality [60]. IUGR means that the fetus cannot achieve the potential size that was genetically predetermined, which is easy to be confused with SGA. Different from SGA, IUGR identifies fetus with risk of poor outcomes without including SGA fetuses that are not pathologically small [61]. IUGR is caused by various reasons including maternal smoking, genetic disorders and infections and severe IUGR cases are associated with impaired placentation [62].

Preterm birth is birth that occurs after 20 weeks' and before 37 weeks' gestation and intensive care is needed for the neonates. Since the lung and brain of a fetus are still developing in the last several weeks of gestation, PTB infants are at great risk of many health issues such as respiratory disorders [63]. Different causes of spontaneous PTB include intrauterine inflammation and placental abnormalities [64]. Different from PTB, miscarriage means a pregnancy loss before 20 weeks' due to natural causes. Main genetic causes of miscarriage are abnormal chromosomes [65].

Gestational diabetes mellitus is defined as glucose intolerance which is diagnosed at 24 to 28 weeks' of pregnancy [66, 67]. As a metabolic disorder, GDM is mainly caused by insulin resistance which increases with gestation and glucose production [68]. GDM increases risk for hypertension, preeclampsia and large infant that increases the probability of caesarean delivery. Although the pathology of GDM is not fully understood, there are some contributing factors identified such as genetic predisposition and elevated placental growth hormone and progesterone levels [69, 70].

1.2.2 Differences of male and female bearing pregnancies

Fetal sex can influence the pregnancy outcomes and may have long term impact on maternal and fetal health [71]. Sex differences exist during fetal development and influence neonatal morbidity and mortality. Still birth and neonatal deaths have proved to cause greater mortality for males than females [72]. Theoretically, the ratio of the sex of the born fetuses is 1:1, however, studies have shown that actually more male fetuses were delivered than female fetuses [73]. Moreover, the sex-specific placenta functions tend to mediate the sex differences during fetal development [74].

Sex chromosomes are the main contributor to sex differences in the placenta. X chromosome inactivation (XCI) exists in the female human placenta [75] and some X-linked genes were partially expressed in chorionic villi but not other tissues, which is associated with the hypomethylation of the related CpG islands [76]. XCI is reversible in placenta and many reprogramming settings, which is named X chromosome reactivation (XCR) [77]. Global XCR can be induced in cells from placental villi *in vitro* with demethylating agents, which is associated with DNA hypomethylation and instable status of XCR of placenta [78].

1.2.3 Maternal blood

Maternal blood during pregnancy can reflect both maternal and fetal health. Biomarkers identified from blood can be diagnostic tools to assess pregnancy health. Serum protein expression can be used as biomarkers for PE [79]. DNA sequencing of fetal cell free DNA in maternal serum can help understand the status of the fetus,

especially for the identification of trisomies 21, 18 and 13 with a high level of accuracy, aiding in disease management and counselling [80]. Recent large-scale analyses ("omics") allow researchers to identify novel biomarkers (as shown in Figure 1.5) for cancers and other diseases [81]. Integration of omics profiles from the placenta and matched maternal blood across gestation may provide insights into ways to identify women and fetuses at risk early during pregnancy [82]. Indeed, maternal blood may reflect the developmental stage of placenta and fetus and may contain placental derived biomarkers. However, few studies to date have matched placental and maternal blood samples that cover continuous gestational stages [3, 83]. Studies that focus on a comprehensive DNA methylation profile of maternal blood that may reflect the developmental stage of placenta and fetus are necessary.

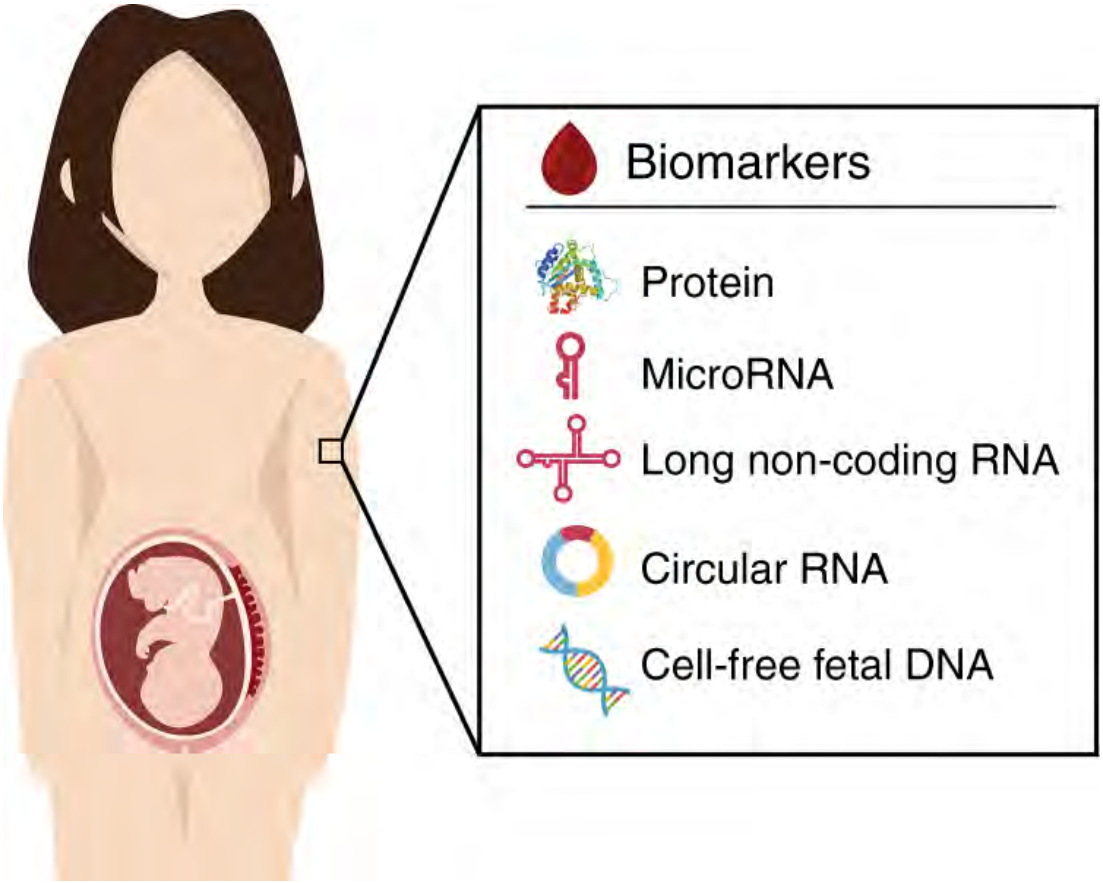

Figure 1.5 Biomarkers in maternal blood that can be used for analysing a fetal disease [81].

1.3 Epigenetic modifications

The term “epigenetics” was first created to describe the processes that are not fully understood during the development of the embryo into a complex organism [84]. While the definition of epigenetics changed with the increased understanding of mechanisms of gene expression and that the same DNA was found in all cells in a specific organism. At present, epigenetics means change of phenotypes caused by chemical modifications of DNA, structural or regulatory factors bound to it without any changes to the underlying DNA sequence [85, 86].

The importance of epigenetic control has long been recognized, but the epigenetic features of some tissue types such as the placenta are still not thoroughly characterised. Most of the known epigenetic modifications such as DNA methylation are reversible which shows the adaptive nature of epigenetic control [87]. Epigenetic modifications stabilise gene expression programmes and may regulate patterns of gene expression [88]. The epigenome can be different between individuals, which highlights its potential in diagnosing disease and evaluating clinical interventions. Here, we describe the major type of epigenetic modifications, DNA methylation, and summarise what is known about their profile within the placenta.

1.3.1 DNA methylation

DNA methylation, the adding of a methyl group, predominantly at the 5^{th} atom of the 6-atom cytosine ring in mammals, is the most widely studied epigenetic modification [89]. An increase in DNA methylation is generally associated with the repression of

gene expression [90]. Methylation can occur at any cytosine base, but it is predominantly found in mammals in the sequence context CpG (mCG) [91]. Recent studies have demonstrated that non-CG methylation (mCH) is correlated with tissue-specific functions [92], but the true biological function of mCH in humans is still unclear and remains to be determined.

In human tissue, DNA methylation is a dynamic process including DNA methylation and demethylation [93]. During DNA methylation, cytosine is transformed to 5'methylcytosine (5mC) by DNA methyltransferases (DNMTs) with the methyl donor S-adenosyl-L-methionine (SAM) [94]. DNA methylation modes fall into either maintenance methylation or *de novo* methylation (Figure 1.6A). While major maintenance methyltransferase DNMT1 maintains the pre-existing methylation patterns, DNMT3A and DNMT3B are *de novo* methyltransferases [95].

Passive demethylation and active demethylation are two DNA demethylation mechanisms, (Figure 1.6B). In passive demethylation, DNA is amplified in an environment without DNMT1, so the DNA is passively demethylated [96]. Enzymes are needed for catalysing or transforming methyl groups during active demethylation. Proteins from the ten-eleven translocation (TET) family have enzymatic activity and conduct active demethylation in mammals [97]. TET hydroxylated 5mC to 5-hydroxymethylcytosine (5hmC) which later can be oxidised by TET again to form 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) and finally lead to the replacement of 5mC with C [98].

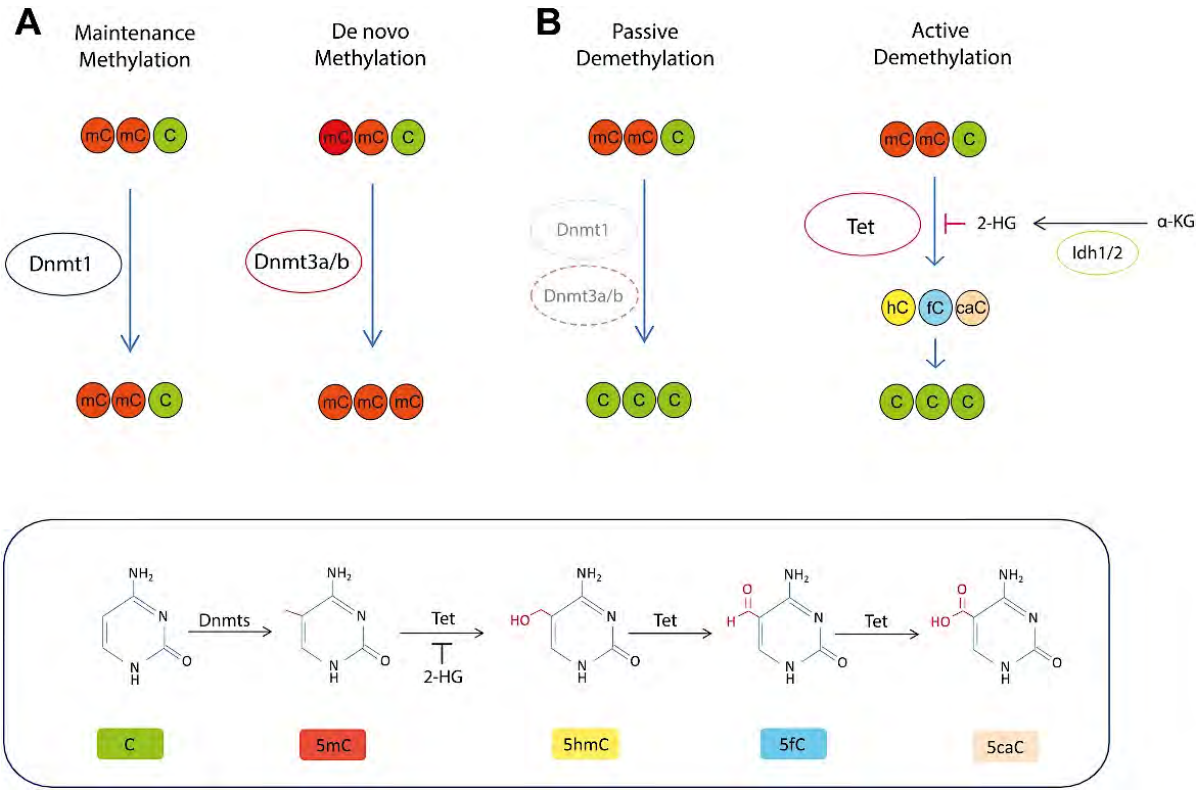

Figure 1.6 The DNA methylation and demethylation processes. (A) DNA methylation process. (B) DNA demethylation process [99].

1.3.2 Epigenetic gene regulation

Many of the studies on gene regulation in chorionic villi focus on gene and transcription level such as epigenetic modifications and transcription factor networks. DNA methylation affect gene expression by influencing the accessibility of regulatory elements [100]. The states of chromatin (“open” or “closed”) were flagged by these epigenetic marks, so they may be maintained during cell replication [101].

DNA methylation can function in single site or CpG clusters in the genome. The CpG clusters are also known as CpG islands with high CG content [102]. Majority of DNA methylation sites that are not in CpG islands such as exons of genes and transposons

are methylated in mammalian cells. DNA methylation sites in CpG islands are usually unmethylated which could be associated with gene promoters, tissue-specific enhancers [103] and some histone modifications that inhibit the binding of *de novo* methylation complex [104].

Canonically, decreased DNA methylation is associated with activation of gene expression and increase of DNA methylation is associated with the repression of gene expression (Figure 1.7). The unmethylated CpG islands are usually associated with activation of genes. Although there is research showing that unmethylated CpG islands are evolutionarily conserved [105], the mechanism of how CpG islands remain unmethylated is not fully understood at present. Increased DNA methylation can cause the condensed state of chromosomes that repress transcription through two mechanisms including recruiting transcriptional repressors that bind to the methylated regions or stop the binding of transcriptional factors to their motifs [100].

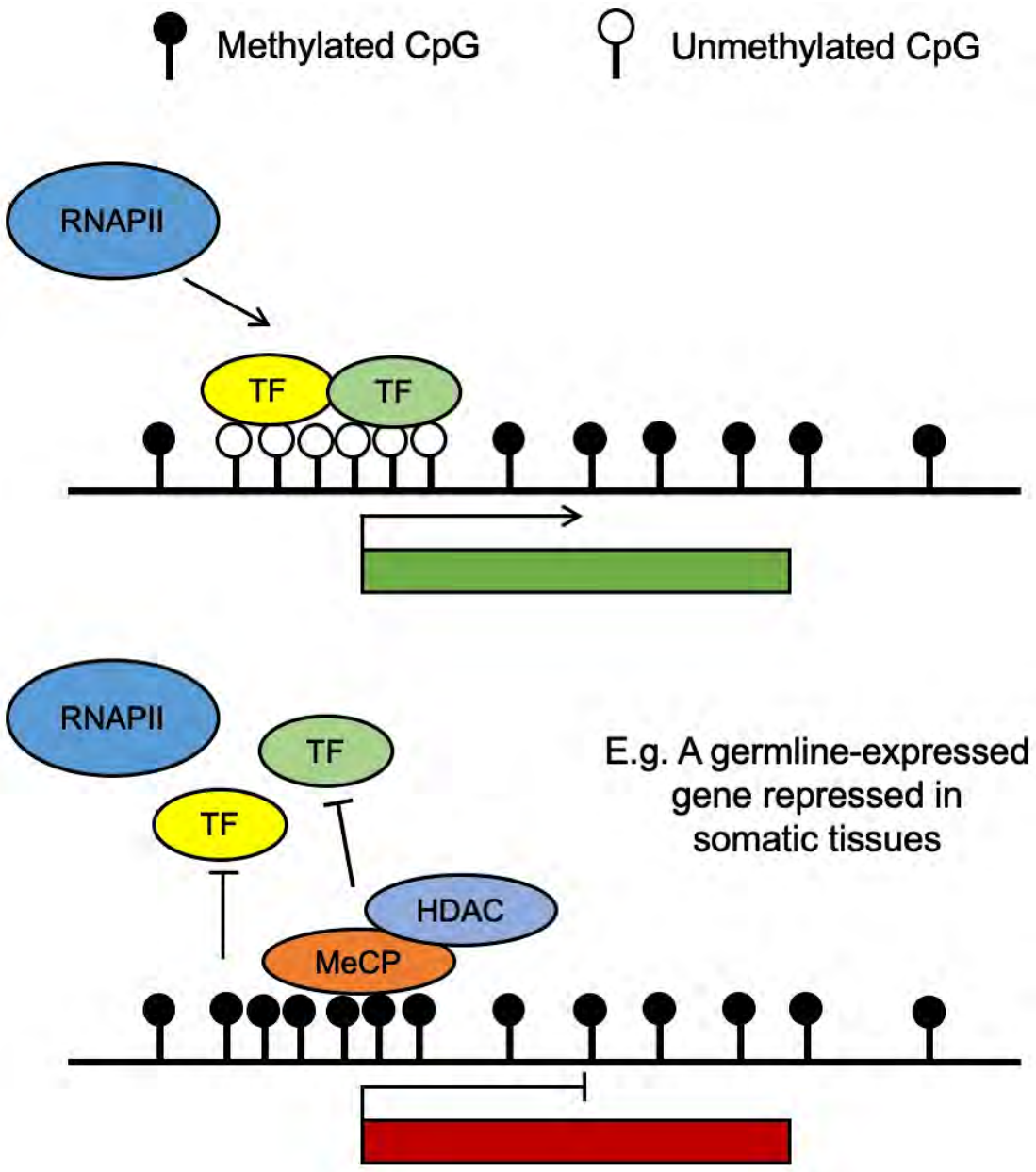

Figure 1.7 The canonical functions of DNA methylation in gene regulation (modified from [106]).

1.3.3 DNA methylation in placenta

The methylome of human placenta is different from most other healthy human tissue types mainly because of the large amount of hypomethylated DNA [9]. However, the origins and roles of hypomethylation in the placenta remain unclear. Since there are

different cell populations that have different methylomes in the placenta, the hypomethylation is likely to reflect the heterogeneity of placental tissue [107, 108]. Another possible reason for DNA hypomethylation is the origin of the placenta. During the preimplantation period of embryo development, the placenta is developed from trophectoderm which is still in a demethylated and pluripotent state [109]. Placental hypomethylation may also be explained by the existence of partially methylated domains (PMDs). Genes in PMDs tend to have function unrelated to the tissue of origin and are repressed [9]. The term placenta is characterised by (PMDs) that cover about 37% of the placental genome and are only found in tissues such as placenta, cancer tissues and immortalised cell lines [108] [9]. In addition, studies showed that DNMT1 was down regulated in trophoblasts which may potentially reduce the global methylation level of placenta [110].

There are many factors including fetal sex, gestational age, and diseases that influence DNA methylation in placenta [111, 112]. Based on DNA methylation microarray data of placenta samples from different trimesters across gestation, Novakovic *et al*. showed that the overall DNA methylation of placenta increased across gestation [11], which was confirmed by subsequent research [15, 113]. Diseases such as PE and GDM are associated with placental epigenetic changes including DNA methylation. Mainly term placental tissue from PE, IUGR and GDM patients were studied and profiled and differentially methylated loci associated with PE, IUGR and GDM were identified using microarray and sequencing technologies [114-120].

1.3.4 Imprinting in placenta

Placenta has a unique epigenetic profile compared with other human somatic tissues. Placenta imprinted genes are also very important and known to regulate placental development and placentation [121]. Gene imprinting is a kind of epigenetic modification that enables the gene to express in a maternal or paternal specific manner which is a potential mechanism of evolution to balance the parental resource allocation to the child in mammals [122]. In general, the paternally expressed genes and maternally expressed genes are associated with accelerated and reduced growth of fetuses and placentas respectively. This phenomenon is consistent with the assumption that the function of imprinting is to deal with the conflict of parental resource allocation to the growth of offspring [123].

Currently, conflicting evidence about placenta imprinted genes exists. Studies have shown that there are 75 human imprinted genes in placenta and 27 of them are placenta-specific [124]. Conversely, in a very recent study using RNA sequencing data of 54 human placenta samples, 50% of the candidate imprinted genes were not expressed or lowly expressed in placenta and only 11 genes were paternally or maternally expressed, including *H19*, *MEG3*, *PHLDA2*, *RTL1*, *AIM1*, *DLK1*, *IGF2*, *MEST*, *PEG10*, *PLAGL1* and *ZFAT*, which indicates an overestimation of the number of imprinted genes in placenta [125]. Besides, new imprinted genes in the placenta are continuously identified, such as *DSCAM* gene [126]. Further studies focused on confirming placental imprinted genes and their functions are still needed.

1.4 High-throughput methods to detect DNA methylation

Genome-wide profiling by microarray- and sequencing-based methods, with their increasingly wide availability, decreasing cost, and continuously improving technology, are powerful tools to map DNA methylation. The gold standard for studying DNA methylation is whole-genome bisulfite sequencing (WGBS). In most human tissue types, except for the regulatory regions that have low methylated sites, the majority (70-80%) of the DNA methylation sites were methylated [127]. Currently, WGBS is still expensive and computationally intensive when there are lots of samples in a study [127].

The Illumina Infinium® Methylation450K BeadChips (450K arrays) and Illumina Infinium® MethylationEPIC BeadChips (EPIC arrays) are widely used platforms for detecting DNA methylation for a large number of samples [128]. The HumanMethylationEPIC (EPIC) array was released in 2015 [129], with almost twice the number of DNA methylation sites of 450K arrays. The increased resolution and expanded coverage of regulatory regions of the human genome, as well as its reasonable pricing, make the EPIC array a suitable and attractive platform for performing profiling of DNA methylation in a large number of samples [129].

Two types of assays exist for detecting DNA methylation on the Infinium methylation array, they are type I and type II assays (Figure 1.8). The type I assay mainly detects DNA methylation sites located in regions of high CpG density and the type II assay mainly focuses on the DNA methylation sites located at low CpG density and mCH [130]. For type I assay, two different query probes, one “unmethylated” and one “methylated”, are used for detecting methylation and unmethylation of one locus

respectively. While for the type II assay, only one query probe is designed for each locus, the methylation status of one site is determined by the single base extension (SBE) result. The complementary nucleotide to the ‘methylated’ C reflects the methylation of one site, and the complementary nucleotide to the ‘unmethylated’ T showed contrary result [131].

The major considerations when choosing between microarray and bisulfite sequencing approaches are cost-effectiveness and data analysis. Commercially available microarrays enable the analysis of CpG methylation status in thousands of gene promoters with lower expenses. WGBS provides a diverse toolbox for DNA methylation analysis with different throughput levels and sequencing depths (i.e. number of times a nucleotide is read).

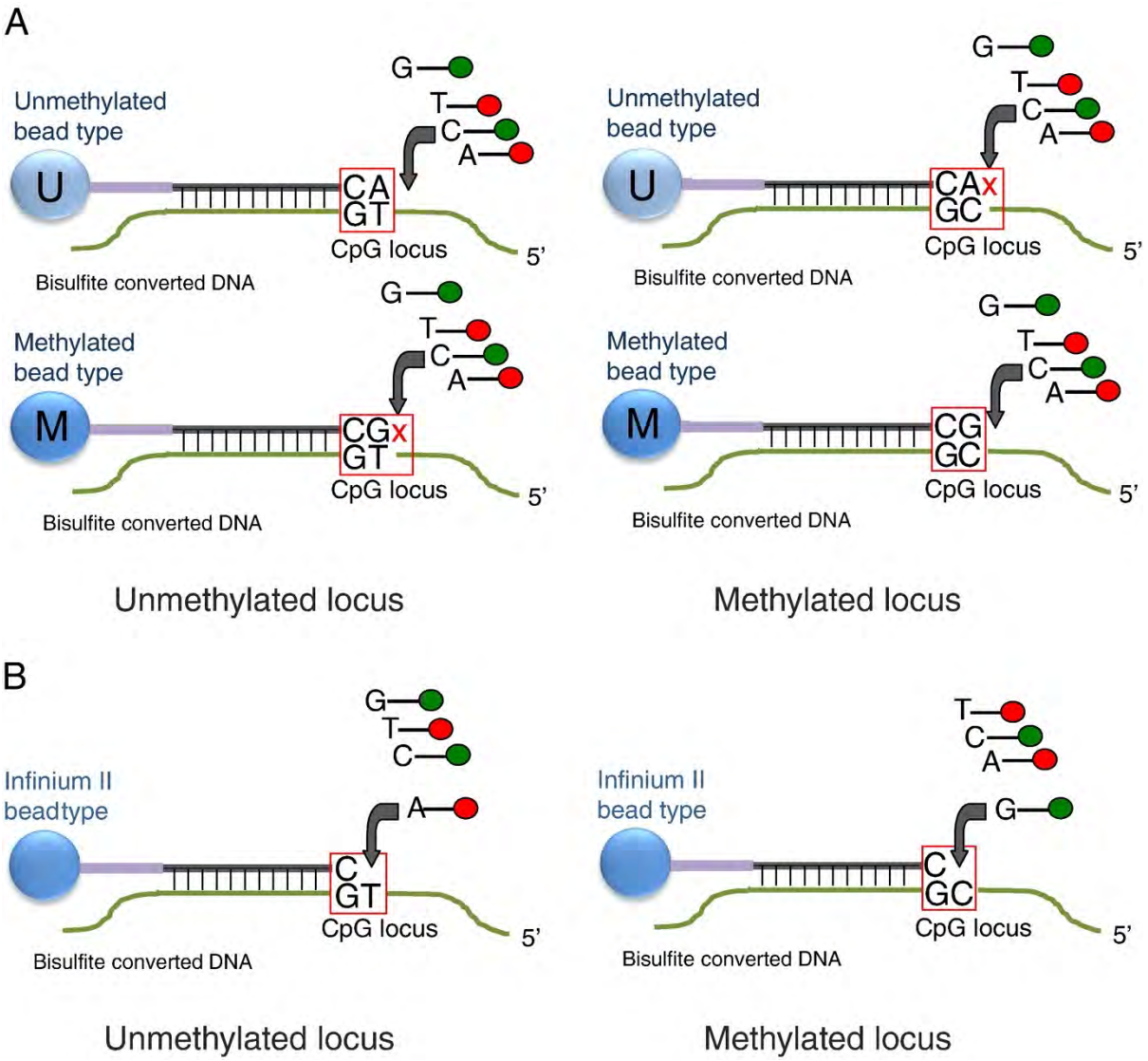

Figure 1.8 Detection of DNA methylation. (A) Infinium I methylation assay scheme. (B) Infinium II methylation assay scheme [132].

1.5 Bioinformatics methods for DNA methylation array data

The analytic pipeline is the foundation for ensuring the reliability and reproducibility of results [133]. There are several published pipelines for DNA methylation array data, and they include *minfi* [134], *RnBeads* [135] and *ChAMP* [136] can process data from EPIC array. Maksimovic et al. also published a cross-package pipeline to analyse DNA methylation array data [137]. However, there is not any established pipeline focusing

on analysing the DNA methylation profile of samples across different time points comprehensively. In order to establish proper pipeline focusing on analysing the DNA methylation profile we reviewed methods for detecting differentially methylated positions (DMPs) and differentially methylated regions (DMRs) as identifying DMPs and DMRs are important steps for DNA methylation analytic pipelines.

1.5.1 Differential methylation

Differential methylation analysis is the process of comparing methylation profiles from two conditions and identifying positions or regions that contain significant changes. Identification of DMP is central in most epigenetic studies [138]. Beta values or M values are commonly used to analyse differential methylation with Beta values being the proportion of methylation intensity over the total intensity (adding intensities from green and red channels). M values on the other hand are logit transformed beta values ($\mathrm{M} = \log 2\left(\frac{\beta}{1-\beta}\right)$) that better represent variance within the data than beta values and previous research has shown that M values performed better than beta values (proportions of methylation intensity) in fitting the Gaussian model [139, 140]. Non-linear methods do exist, such as using beta regression models (*gamlss* package) [141], however linear model methods contained in packages *limma* [142], *missMethyl* [143] and *ruv* [144] are far more popular.

There are advantages and disadvantages for both linear and non-linear models. The main advantage of the non-linear beta regression is that it directly uses beta values that are biologically meaningful. Non-linear models allow dependent observations while linear models assume that observations are independent from each other.

However, linear models are more widely used for analysing microarray data and implemented in different pipeline and packages. The outcomes from analysing M values are also generally visualised with Beta values [145]. In Chapter 2, we compared different models for identifying significantly changed probes between groups and found that multivariant linear model using M values is the method for identifying changes of DNA methylation between groups, since it is also used in methods of identifying differently methylated regions and the results can be interpreted properly. Linear regression is considered as a valid statistical method for analysing DNA methylation array data, although M values do not always meet the assumption of regression tests [146].

DMR are genomic regions in which DNA methylation is consistently positively or negatively associated with a phenotype or exposure and the identification of DMRs is very important for understating the etiology of diseases [147]. Compared to methods for identifying DMP, there are fewer methods for identifying DMR because for identifying DMP the methods for identifying differentially expressed genes can be used for reference. The package *bumphunter* separates probes into clusters based on their positions in the genome, with beta values of one probe in a specific position are fitted in a linear model to estimate the loess smooth function with each point weighted. Statistical uncertainty of each estimated DMR are tested using bootstrap or permutation tests [148]. Family wise error rate (FWER) [149] for each DMR represents the significance of the DMR by comparing permutated DMR to the observed DMR. However, by using the loess method, *bumphunter* filters out regions with sparse site density, so the result form *bumphunter* is affected by false negatives [150].

Conversely to the Probe Lasso method implemented in *ChAMP* package [136], M values are first fitted into a linear model using *limma* package for differentially methylation analysis at the probe level. Flexible windows for identifying DMR are then generated according to genetic and epigenetic annotations of each probe of the methylation array. Differentially methylated probes fall into 28 genetic/epigenetic categories including 7 gene features: 1500bp of a transcription start site (TSS1500), 200bp of a transcription start site (TSS200), 5 prime untranslated regions (5'UTR), first exons of genes (1stExon), gene body (Body), 3 prime untranslated regions (3'UTR) and intergenic regions (IGR), and 4 CGI relations (Island, Shore, Shelf and Open sea). Dynamic window/lasso sizes (28 windows) are assigned to each genetic/epigenetic category based on the difference of probe spacing features between these categories [151]. For example, within 200bp of a transcription start site (TSS200), the probe density is higher than probe density in 3'UTR regions, so smaller windows (7~24bp) are used for TSS200, whereas probes on 3'UTR regions are combined in a bigger window (69~1949bp) [151]. If the gap between two DMR is less than a specific use defined threshold (normally 1000bp, since DNA methylation sites within 1000 bp have been shown to be significantly correlated [152]), the two DMR were merged. Stouffer's method is then used to combine and estimate the final p-values for DMR which are then adjusted for multiple tests.

For DMRcate method, standard linear models (use *limma*) are first applied to M values to analyse differential methylation at the probe level [150]. The signed t statistics between two groups from *limma* for each CpG cite are transferred to unsigned F statistics which is used in the following steps. Gaussian smoothing instead of loess in *bumphunter* is used for weighing and smoothing the per-CpG-site test statistics. The

Gaussian smoothed test statistics are then modelled by a scaled chi-squared random variable. For each chromosome, two smoothed estimates are computed: one weighted and one not, for a null comparison. The two estimates are compared via a Satterthwaite approximation [153] and a significance test is calculated at all hg19 coordinates where the probes map to. Finally, nearby significant CpG sites that are at most λ bp (usually λ=1000) between each other and had FDR-corrected p-values smaller than 0.05 (cutoff of p-values was automatically determined by the number of significant CpGs returned by *limma*; a small p-value means the probe is in a location surrounded by many other probes) are agglomerated as DMR [150]. For an overall significance metric per DMR, Stouffer's method is used to combine FDRs derived from *limma* to estimate p-values for each DMR.

There are advantages and disadvantages of these methods for detecting DMR. *Bumphunter* normally has less overlapping results with other methods [150, 154]. *Bumphunter* mainly detects DMR with methylation changes in the same direction (either hyper- or hypo-methylated) and the DMR should be long enough to be identified by loess smoother, which lead to false negative results [155]. Probe Lasso method considers the biological features of DNA which is the advantage of this method [151], while the drawback is that the true DMR that have multiple annotations can be split and consequently reducing the sensitivity of the method [150]. Compared to *Bumphunter* and *Probe Lasso*, DMRcate is more sensitive to DMR, while the limitations of DMRcate are: DMRcate could only compare 2 categorical variables (e.g. case & control). In Chapter 2, we recommend that using different methods with the same data set to detect DMR because they can be complementary to one another, which is also proved in previous study [147].

1.5.2 Partially methylated domains and imprinted genes

As previously mentioned above, the placenta is unique in that it contains partially methylated domains, which are characterised as large hypomethylated regions in the genome. At present, there is no method developed for identifying PMDs using EPIC or 450K array data since array data has low coverage compared to data from WGBS [156]. There are several methods to identify PMDs with whole genome bisulfite sequencing (WGBS) including using a sliding window approach to identify PMDs [157, 158], using two states Hidden Markov Model (HMM) to identify PMDs [159, 160] and recently, the random forest classifier is also used to identify PMDs [161].

Imprinted genes are very important for the function and normal development of placenta [162]. Imprinted genes and imprinting control regions are identified using different methods by different research groups. Using 450K array data from diandric (homologous) and digynic triploid (containing more than two paired sets of chromosomes) placenta samples, imprinted DMRs were identified [163]. Imprinting control regions are also identified by DMRs between samples from imprinting disorders and healthy controls [164]. All the imprinted genes identified using single-cell sequencing data and imprinting control regions (ICRs) identified using WGBS data [165, 166] are listed at the website: http://www.geneimprint.com/site/genes-by-species.

1.5.3 Quality control of array data and placenta tissue samples

There are methods to detecting the failed samples and probes caused by the variation during library preparation and array processing. For example, the sample-dependent control probes in array were used to check the quality of bisulfite conversion for type I and II assays, non-polymorphic performance, specificity of matching type I and II probes and the system background [134]. Undetected probes and failed samples can be filtered out using the cut-offs for detection *p* values [167].

Comparing to quality control of array processing steps, there are less methods for quality control of samples. The samples that mixed with other tissue types can influence the accuracy of DNA methylation analyses. Previous studies take advantage of the DNA methylation of SNP sites to estimate the purity of samples and highlighted the importance of assessing sample purity as a quality control step [168]. Whether this method is suitable to identify mixed placenta samples was investigated in Chapter 3, which is published in *Placenta* [169].

1.6 Application of DNA methylation analyses

1.6.1 Biomarkers

Biomarkers are measurements of the association and interaction between biological system and potential adverse factors such as poisons. These measurements can reflect the changes of tissue functions or physiology or changes at cell and molecular level [170]. A lot of features can be used as biomarkers including blood pressure, blood

and saliva, these biomarkers were used to predict clinical outcomes, environmental exposure and future health. The most important issue for now is confirming the association between the biomarker and the clinical outcome.

Biomarkers from placenta include RNA-based, DNA-based and protein-based biomarkers, could be used for diagnosing and predicting some pregnancy complications such as PE and PTB [171, 172]. The profiling of transcriptome enables the improvement of clinical diagnoses. For example, by profiling the gene expression of placenta from normal and complicated pregnancies, Leavey et al. identified different placenta phenotypes and 3 of them are related to PE [173]. However, the limitation of using RNA as biomarkers for pregnancy outcomes is that the placenta is difficult to sample in an ongoing pregnancy, hence the need to find biomarkers in maternal blood that reflect placental function [174]. Since it is hard to sample placenta, the non-invasive tests using blood samples or other bodily fluids are more popular. The proteins in maternal serum including pregnancy associated plasma protein-A (PAPPA), placental growth factor (PIGF), human chorionic gonadotropin (hCG), inhibin Subunit Alpha (INHA) and soluble fms-like tyrosine kinase-1 (sFLT1) are widely used to predict poor pregnancy outcomes [175]. Fetal cell-free DNA in maternal blood is derived from the placental syncytiotrophoblast, and it is a DNA based biomarker that is very accurate in screening trisomies [176].

DNA methylation of placenta, maternal blood (including plasma or buffy coat) can be used for predicting future child health or the intrauterine environment for early intervention. Studies have shown that the adverse influence on epigenome from early fetal environment can be prevented or reversed [177]. Through the epigenetic

biomarkers from early pregnancy, it is possible to not only identify pregnancy complications that potentially develop into severe disease in early gestation but also develop Intervention strategies that can target and reverse the epigenetic changes [177].

1.6.2 Prediction of gestational age, sex, ethnicity, environmental exposure, disease and gene expression

Studies so far indicate that DNA methylation profile could be used to predict age, gestational age (GA), sex, ethnicity, environmental exposure, disease and gene expression [112, 178]. Using DNA methylation array datasets data sets of 8000 samples from 51 healthy tissues and cell types, Horvath et al. identified 353 DNA methylation sites (probes selected from common probes between 450k and 27k array platform) to form the epigenetic clock using elastic net regression model. This epigenetic clock can be used to predict age and also estimated the potential altered aging process in disease [179]. It is important to notice that Horvath epigenetic clock is established using probes in common between 450K and 27K array, it may not be the fully optimised epigenetic clock for EPIC arrays [180].

Gestational age (GA) is used as a sign for fetal development by researchers and clinicians. Previous study showed that DNA methylation of neonatal cord blood and placenta can be used for predicting GA [181]. DNA methylation profile at birth were predictors of GA [182]. However, epigenetic clocks derived from cord blood or other tissues do not accurately estimate GA in placental samples. Studies also showed that

DNA methylation of placenta across gestation can be used to predict GA accurately [15, 183].

DNA methylation pattern of sex chromosomes of women is special because of the X chromosome inactivation [184]. Since DNA methylation in sex chromosomes is dramatically different between female and male, it can accurately predict the sex of samples [185]. Methods that estimate sex are based on the intensity measures of DNA methylation sites on the X and Y chromosomes respectively [134]. Although the differences of DNA methylation between female and male samples mainly exist on the X chromosome, it has also been shown that sex also influences the methylation of autosomal genes [186].

Recently, the DNA methylation profile is used to predict ethnicity for placenta samples. Yuan et al. developed the ethnicity classifier based on elastic net model which can be used for predicting ethnicity when the meta data of the public data is not complete. The ethnicity predictor is developed based on DNA methylation sites on 450K arrays and about 50% of the sites consist of the predictor are associated with genetic polymorphisms [187].

DNA methylation is also can be used to predict environmental exposures [112]. For example, air pollution is correlated with increased DNA methylation of placenta [188]. Maternal smoking also may alter DNA methylation of placenta and maternal blood [189]. As the burden of environmental exposures such as air and water pollution increased, altered DNA methylation caused by pollutants can be good indicators.

Moreover, DNA methylation profile is also used to predict gene expression. As DNA methylation especially the change of DNA methylation at gene promoters were usually negatively correlated with gene expression. Instead of elastic net regression that is usually used for predicting age and ethnicity, the predictor of gene expression is established using least absolute shrinkage and selection operator (LASSO) regression, however, the power of this predictor may change with the changing of tissue or cell types [190].

1.7 Summary

The overall aim of this thesis is to use bioinformatic methods to establish an analytic pipeline and profiling DNA methylation in placenta from terminated pregnancies of early gestation and the corresponding maternal leukocytes. Adopting suitable methods in the profiling pipeline will allow us to have results that can be interpreted properly. The quality control method of detecting placenta tissue purity was established and improved in this study. The analyses focusing on DNA methylation data from EPIC arrays and the corresponding RNA sequencing data to understand the regulatory functions of DNA methylation changes in placenta. The DNA methylation profile of matched maternal leukocytes were also investigated aiming at finding relevant biomarkers associated with the changing of phenotypes of placenta and maternal status.

References

1. Barker DJ. The fetal and infant origins of adult disease. BMJ: British Medical Journal. 1990;301(6761):1111.
2. Moore KL, Persaud TVN, Torchia MG. The Developing Human: Elsevier Health Sciences; 2011.
3. Burton GJ, Fowden AL, Thornburg KL. Placental origins of chronic disease. Physiological reviews. 2016;96(4):1509-65.
4. Burton GJ, Kaufmann P, Huppertz B. Anatomy and genesis of the placenta. Knobil and Neill's physiology of reproduction. 2006;3:189-243.
5. Brett KE, Ferraro ZM, Yockell-Lelievre J, Gruslin A, Adamo KB. Maternal–fetal nutrient transport in pregnancy pathologies: the role of the placenta. International journal of molecular sciences. 2014;15(9):16153-85.
6. O'Tierney-Ginn P, Presley L, Myers S, Catalano P. Placental growth response to maternal insulin in early pregnancy. The Journal of Clinical Endocrinology & Metabolism. 2015;100(1):159-65.
7. Geldenhuys J, Rossouw TM, Lombaard HA, Ehlers MM, Kock MM. Disruption in the Regulation of immune Responses in the Placental Subtype of Preeclampsia. Frontiers in immunology. 2018;9.
8. Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res. 1982 Apr 24;10(8):2709-21. DOI: 10.1093/nar/10.8.2709.
9. Schroeder DI, Blair JD, Lott P, Yu HOK, Hong D, Crary F, et al. The human placenta methylome. Proceedings of the national academy of sciences. 2013;110(15):6037-42.

10. Novakovic B, Saffery R. Placental pseudo-malignancy from a DNA methylation perspective: unanswered questions and future directions. Frontiers in genetics. 2013;4.

11. Novakovic B, Yuen RK, Gordon L, Penaherrera MS, Sharkey A, Moffett A, et al. Evidence for widespread changes in promoter methylation profile in human placenta in response to increasing gestational age and environmental/stochastic factors. BMC genomics. 2011;12(1):529.

12. He Z, Lu H, Luo H, Gao F, Wang T, Gao Y, et al. The promoter methylomes of monochorionic twin placentas reveal intrauterine growth restriction-specific variations in the methylation patterns. Scientific reports. 2016;6.

13. Rouault C, Clément K, Guesnon M, Henegar C, Charles M-A, Heude B, et al. Transcriptomic signatures of villous cytotrophoblast and syncytiotrophoblast in term human placenta. Placenta. 2016;44:83-90.

14. Haider S, Meinhardt G, Saleh L, Fiala C, Pollheimer J, Knöfler M. Notch1 controls development of the extravillous trophoblast lineage in the human placenta. Proceedings of the National Academy of Sciences. 2016:201612335.

15. Mayne BT, Leemaqz SY, Smith AK, Breen J, Roberts CT, Bianco-Miotto T. Accelerated placental aging in early onset preeclampsia pregnancies identified by DNA methylation. Epigenomics. 2017 Mar;9(3):279-89. DOI: 10.2217/epi-2016-0103.

16. Bartlett A, Lewis J, Williams ML. Generations of interdisciplinarity in bioinformatics. New Genet Soc. 2016 Apr 02;35(2):186-209. DOI: 10.1080/14636778.2016.1184965.

17. Turco MY, Moffett A. Development of the human placenta. Development. 2019;146(22).

18. Turco MY, Moffett A. Development of the human placenta. Development. 2019;146(22):dev163428. DOI: 10.1242/dev.163428.

19. Benirschke K, Burton GJ, Baergen RN. Early development of the human placenta. Pathology of the human placenta: Springer; 2012. p. 41-53.

20. Li Z, Kurosawa O, Iwata H. Establishment of human trophoblast stem cells from human induced pluripotent stem cell-derived cystic cells under micromesh culture. Stem cell research & therapy. 2019;10(1):1-14.

21. Gude NM, Roberts CT, Kalionis B, King RG. Growth and function of the normal human placenta. Thromb Res. 2004;114(5-6):397-407. DOI: 10.1016/j.thromres.2004.06.038.

22. Huppertz B. The anatomy of the normal placenta. Journal of clinical pathology. 2008;61(12):1296-302.

23. Frendo JL, Vidaud M, Guibourdenche J, Luton D, Muller F, Bellet D, et al. Defect of Villous Cytotrophoblast Differentiation into Syncytiotrophoblast in Down's Syndrome1. The Journal of Clinical Endocrinology & Metabolism. 2000;85(10):3700-7. DOI: 10.1210/jcem.85.10.6915.

24. Fogarty N, Mayhew T, Ferguson-Smith A, Burton G. A quantitative analysis of transcriptionally active syncytiotrophoblast nuclei across human gestation. Journal of anatomy. 2011;219(5):601-10.

25. Pfeffer PL, Pearton DJ. Trophoblast development. Reproduction (Cambridge, England). 2012;143(3):231-46.

26. Donnelly L, Campling G. Functions of the placenta. Anaesthesia & intensive care medicine. 2014;15(3):136-9.

27. Zygmunt M, Herr F, Münstedt K, Lang U, Liang OD. Angiogenesis and vasculogenesis in pregnancy. European Journal of Obstetrics & Gynecology and Reproductive Biology. 2003;110:S10-S8.

28. Moore KL, Persaud TVN, Torchia MG. Before We Are Born E-Book: Essentials of Embryology and Birth Defects With STUDENT CONSULT Online Access: Elsevier Health Sciences; 2015.

29. Wang Y. Vascular biology of the placenta. In: Colloquium Series on Integrated Systems Physiology: from Molecule to Function. 2010: Morgan & Claypool Life Sciences; p. 1-98.

30. Burton GJ, Jauniaux E. The cytotrophoblastic shell and complications of pregnancy. Placenta. 2017;60:134-9.

31. Rasmussen KM, Yaktine AL. Composition and components of gestational weight gain: physiology and metabolism. Weight Gain During Pregnancy: Reexamining the Guidelines: National Academies Press (US); 2009.

32. Burton GJ, Kaufmann P, Huppertz B. Anatomy and genesis of the placenta. Knobil and Neill's physiology of reproduction. 2006;1:189-244.

33. Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), NIH, HHS. Available from: https://www.nichd.nih.gov/

34. De Miguel M, Alcaina Y, Sainz de la Maza D, Lopez-Iglesias P. Cell metabolism under microenvironmental low oxygen tension levels in stemness, proliferation and pluripotency. Current molecular medicine. 2015;15(4):343-59.

35. Caniggia I, Winter J, Lye SJ, Post M. Oxygen and placental development during the first trimester: implications for the pathophysiology of pre-eclampsia. Placenta. 2000 Mar-Apr;21 Suppl A:S25-30.

36. Burton GJ, Hempstock J, Jauniaux E. Oxygen, early embryonic metabolism and free radical-mediated embryopathies. Reprod Biomed Online. 2003 Jan-Feb;6(1):84-96.

37. Jauniaux E, Watson AL, Hempstock J, Bao YP, Skepper JN, Burton GJ. Onset of maternal arterial blood flow and placental oxidative stress. A possible factor in human early pregnancy failure. Am J Pathol. 2000 Dec;157(6):2111-22. DOI: 10.1016/S0002-9440(10)64849-3.

38. He C, Shan N, Xu P, Ge H, Yuan Y, Liu Y, et al. Hypoxia-induced Downregulation of SRC-3 Suppresses Trophoblastic Invasion and Migration Through Inhibition of the AKT/mTOR Pathway: Implications for the Pathogenesis of Preeclampsia. Scientific reports. 2019;9(1):1-12.

39. Gutierrez JA, Gomez I, Chiarello DI, Salsoso R, Klein AD, Guzman-Gutierrez E, et al. Role of proteases in dysfunctional placental vascular remodelling in preeclampsia. Biochim Biophys Acta Mol Basis Dis. 2019 Apr 5. DOI: 10.1016/j.bbadis.2019.04.004.

40. Pereira RD, De Long NE, Wang RC, Yazdi FT, Holloway AC, Raha S. Angiogenesis in the placenta: the role of reactive oxygen species signaling. Biomed Res Int. 2015;2015:814543. DOI: 10.1155/2015/814543.

41. Carter AM, Enders AC. Comparative aspects of trophoblast development and placentation. Reproductive Biology and Endocrinology. 2004;2(1):46.

42. Guttmacher AE, Maddox YT, Spong CY. The Human Placenta Project: placental structure, development, and function in real time. Placenta. 2014;35(5):303-4.

43. Gaunt G, Ramin K. Immunological tolerance of the human fetus. American journal of perinatology. 2001;18(06):299-312.

44. PrabhuDas M, Bonney E, Caron K, Dey S, Erlebacher A, Fazleabas A, et al. Immune mechanisms at the maternal-fetal interface: perspectives and challenges. Nature immunology. 2015;16(4):328-34.

45. Jansson T. Placenta plays a critical role in maternal-fetal resource allocation. Proc Natl Acad Sci U S A. 2016 Oct 4;113(40):11066-8. DOI: 10.1073/pnas.1613437113.

46. Hsiao EY, Patterson PH. Placental regulation of maternal-fetal interactions and brain development. Developmental neurobiology. 2012;72(10):1317-26.

47. Fan X, Rai A, Kambham N, Sung JF, Singh N, Petitt M, et al. Endometrial VEGF induces placental sFLT1 and leads to pregnancy complications. J Clin Invest. 2014 Nov;124(11):4941-52. DOI: 10.1172/JCI76864.

48. Steegers EA, von Dadelszen P, Duvekot JJ, Pijnenborg R. Pre-eclampsia. The Lancet. 2010;376(9741):631-44.

49. Wang KC, Larcombe A, Morton J, Davidge S, James A, Noble PB. Maternal Hypoxia-Induced Intrauterine Growth Restriction Predisposed Offspring To Airway Hyperresponsiveness In Adulthood. A29 INFLAMMATION AND MECHANISMS OF AIRWAY SMOOTH MUSCLE CONTRACTION: Am Thoracic Soc; 2016. p. A1264-A.

50. Grigsby PL. Animal Models to Study Placental Development and Function throughout Normal and Dysfunctional Human Pregnancy. Semin Reprod Med. 2016 Jan;34(1):11-6. DOI: 10.1055/s-0035-1570031.

51. Lindgren K. Relationships among maternal–fetal attachment, prenatal depression, and health practices in pregnancy. Research in nursing & health. 2001;24(3):203-17.

52. Duley L. The global impact of pre-eclampsia and eclampsia. In: Seminars in perinatology. 2009: Elsevier; p. 130-7.

53. Mol BWJ, Roberts CT, Thangaratinam S, Magee LA, de Groot CJM, Hofmeyr GJ. Pre-eclampsia. Lancet. 2016 Mar 5;387(10022):999-1011. DOI: 10.1016/S0140-6736(15)00070-7.

54. English FA, Kenny LC, McCarthy FP. Risk factors and effective management of preeclampsia. Integrated blood pressure control. 2015;8:7.

55. Grimes S, Bombay K, Lanes A, Walker M, Corsi DJ. Potential biological therapies for severe preeclampsia: a systematic review and meta-analysis. BMC pregnancy and childbirth. 2019;19(1):163.

56. Brown MA, Magee LA, Kenny LC, Karumanchi SA, McCarthy FP, Saito S, et al. Hypertensive disorders of pregnancy: ISSHP classification, diagnosis, and management recommendations for international practice. Hypertension. 2018;72(1):24-43.

57. Mayrink J, Costa M, Cecatti J. Preeclampsia in 2018: revisiting concepts, physiopathology, and prediction. The Scientific World Journal. 2018;2018.

58. Kim YJ. Pathogenesis and promising non-invasive markers for preeclampsia. Obstetrics & gynecology science. 2013;56(1):2-7.

59. Ludvigsson JF, Lu D, Hammarström L, Cnattingius S, Fang F. Small for gestational age and risk of childhood mortality: A Swedish population study. PLoS medicine. 2018;15(12):e1002717.

60. Tachibana M, Nakayama M, Ida S, Kitajima H, Mitsuda N, Ozono K, et al. Pathological examination of the placenta in small for gestational age (SGA) children with or without postnatal catch-up growth. The Journal of Maternal-Fetal & Neonatal Medicine. 2016;29(6):982-6.

61. Ross MG, Mansano RZ. Fetal growth restriction. Update. 2013.

62. Veerbeek J, Nikkels P, Torrance H, Gravesteijn J, Uiterweer EP, Derks J, et al. Placental pathology in early intrauterine growth restriction associated with maternal hypertension. Placenta. 2014;35(9):696-701.

63. Boyle AK, Rinaldi SF, Norman JE, Stock SJ. Preterm birth: Inflammation, fetal injury and treatment strategies. Journal of reproductive immunology. 2017;119:62-6.

64. Keelan JA, Newnham JP. Recent advances in the prevention of preterm birth. F1000Research. 2017;6.

65. Colley E, Hamilton S, Smith P, Morgan NV, Coomarasamy A, Allen S. Potential genetic causes of miscarriage in euploid pregnancies: a systematic review. Human reproduction update. 2019;25(4):452-72.

66. American Diabetes A. 2. Classification and Diagnosis of Diabetes. Diabetes Care. 2016 Jan;39 Suppl 1:S13-22. DOI: 10.2337/dc16-S005.

67. Gilbert L, Gross J, Lanzi S, Quansah DY, Puder J, Horsch A. How diet, physical activity and psychosocial well-being interact in women with gestational diabetes mellitus: an integrative review. BMC Pregnancy Childbirth. 2019 Feb 7;19(1):60. DOI: 10.1186/s12884-019-2185-y.

68. Newbern D, Freemark M. Placental hormones and the control of maternal metabolism and fetal growth. Current Opinion in Endocrinology, Diabetes and Obesity. 2011;18(6):409-16.

69. Agha-Jaffar R, Oliver N, Johnston D, Robinson S. Gestational diabetes mellitus: does an effective prevention strategy exist? Nature Reviews Endocrinology. 2016;12(9):533.

70. Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The pathophysiology of gestational diabetes mellitus. International journal of molecular sciences. 2018;19(11):3342.

71. Al-Qaraghouli M, Fang YMV. Effect of fetal sex on maternal and obstetric outcomes. Frontiers in pediatrics. 2017;5:144.

72. Zhao D, Zou L, Lei X, Zhang Y. Gender differences in infant mortality and neonatal morbidity in mixed-gender twins. Scientific reports. 2017;7(1):1-6.

73. Vatten LJ, Skjærven R. Offspring sex and pregnancy outcome by length of gestation. Early human development. 2004;76(1):47-54.

74. Clifton V. Sex and the human placenta: mediating differential strategies of fetal growth and survival. Placenta. 2010;31:S33-S9.

75. Zeng S-M, Yankowitz J. X-inactivation patterns in human embryonic and extra-embryonic tissues. Placenta. 2003;24(2-3):270-5.

76. Migeon BR, Wolf SF, Axelman J, Kaslow DC, Schmidt M. Incomplete X chromosome dosage compensation in chorionic villi of human placenta. Proceedings of the National Academy of Sciences. 1985;82(10):3390-4.

77. Pasque V, Talon I, Janiszewski A, Chappell J, Vanheer L. Recent advances in understanding the reversal of gene silencing during X chromosome reactivation. Frontiers in cell and developmental biology. 2019;7:169.

78. Migeon BR, Axelman J, Jeppesen P. Differential X reactivation in human placental cells: implications for reversal of X inactivation. The American Journal of Human Genetics. 2005;77(3):355-64.

79. Wong F, Cox B. Proteomics analysis of preeclampsia, a systematic review of maternal and fetal compartments. J Proteomics Bioinform S. 2014;10:2.

80. Snyder HL, Curnow KJ, Bhatt S, Bianchi DW. Follow-up of multiple aneuploidies and single monosomies detected by noninvasive prenatal testing: implications for management and counseling. Prenatal diagnosis. 2016;36(3):203-9.

81. Wagner R, Tse WH, Gosemann J-H, Lacher M, Keijzer R. Prenatal maternal biomarkers for the early diagnosis of congenital malformations: A review. Pediatric research. 2019:1.

82. Roberts C. 430. Omics technologies and the future of preeclampsia research and clinical practice. Pregnancy Hypertension. 2018;13:S9.

83. Karin-Kujundzic V, Sola IM, Predavec N, Potkonjak A, Somen E, Mioc P, et al. Novel Epigenetic Biomarkers in Pregnancy-Related Disorders and Cancers. Cells. 2019;8(11):1459.

84. Waddington CH. Embryology, epigenetics and biogenetics. Nature. 1956;177(4522):1241-.

85. Morris J. Genes, genetics, and epigenetics: a correspondence. Science. 2001;293(5532):1103-5.

86. Felsenfeld G. A brief history of epigenetics. Cold Spring Harbor perspectives in biology. 2014;6(1):a018200.

87. Portela A, Esteller M. Epigenetic modifications and human disease. Nature biotechnology. 2010;28(10):1057-68.

88. Kim S, Kaang B-K. Epigenetic regulation and chromatin remodeling in learning and memory. Experimental & molecular medicine. 2017;49(1):e281.

89. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet. [Perspectives]. 2016 08//print;17(8):487-500. DOI: 10.1038/nrg.2016.59.

90. Anastasiadi D, Esteve-Codina A, Piferrer F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. Epigenetics & chromatin. 2018;11(1):37.

91. Ford EE, Grimmer MR, Stolzenburg S, Bogdanovic O, de Mendoza A, Farnham PJ, et al. Frequent lack of repressive capacity of promoter DNA methylation identified through genome-wide epigenomic manipulation. bioRxiv. 2017:170506.

92. Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, et al. Human body epigenome maps reveal noncanonical DNA methylation variation. Nature. 2015;523(7559):212-6.

93. Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, et al. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science. 2013;339(6118):448-52.

94. Jin B, Robertson KD. DNA methyltransferases, DNA damage repair, and cancer. Epigenetic Alterations in Oncogenesis: Springer; 2013. p. 3-29.

95. Chen Z-x, Riggs AD. DNA methylation and demethylation in mammals. Journal of Biological Chemistry. 2011;286(21):18347-53.

96. Shen L, Inoue A, He J, Liu Y, Lu F, Zhang Y. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. Cell stem cell. 2014;15(4):459-70.

97. Pastor WA, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nature reviews Molecular cell biology. 2013;14(6):341.

98. Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. Nat Rev Mol Cell Biol. 2010 Sep;11(9):607-20. DOI: 10.1038/nrm2950.

99. Jeong M, Goodell MA. New answers to old questions from genome-wide maps of DNA methylation in hematopoietic cells. Experimental Hematology. 2014 2014/08/01/;42(8):609-17. DOI: http://dx.doi.org/10.1016/j.exphem.2014.04.008.

100. Choudhuri S, Cui Y, Klaassen CD. Molecular targets of epigenetic regulation and effectors of environmental influences. Toxicology and applied pharmacology. 2010;245(3):378-93.

101. Fritz A, Gillis N, Gerrard D, Rodriguez P, Hong D, Rose J, et al. Higher order genomic organization and epigenetic control maintain cellular identity and prevent breast cancer. Genes, Chromosomes and Cancer. 2019;58(7):484-99.

102. Fazzari MJ, Greally JM. Epigenomics: beyond CpG islands. Nature Reviews Genetics. 2004;5(6):446.

103. Wise TL, Pravtcheva DD. The undermethylated state of a CpG island region in igf2 transgenes is dependent on the H19 enhancers. Genomics. 1999;60(3):258-71.

104. Straussman R, Nejman D, Roberts D, Steinfeld I, Blum B, Benvenisty N, et al. Developmental programming of CpG island methylation profiles in the human genome. Nature structural & molecular biology. 2009;16(5):564.

105. Long HK, King HW, Patient RK, Odom DT, Klose RJ. Protection of CpG islands from DNA methylation is DNA-encoded and evolutionarily conserved. Nucleic acids research. 2016;44(14):6693-706.

106. Reddington JP, Pennings S, Meehan RR. Non-canonical functions of the DNA methylome in gene regulation. Biochemical Journal. 2013;451(1):13-23.

107. Fogarty NM, Burton G, Ferguson-Smith A. Different epigenetic states define syncytiotrophoblast and cytotrophoblast nuclei in the trophoblast of the human placenta. Placenta. 2015;36(8):796-802.

108. Robinson WP, Price EM. The human placental methylome. Cold Spring Harbor perspectives in medicine. 2015;5(5):a023044.

109. Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. Human molecular genetics. 2005;14(suppl_1):R47-R58.

110. Novakovic B, Wong NC, Sibson M, Ng HK, Morley R, Manuelpillai U, et al. DNA methylation-mediated down-regulation of DNA methyltransferase-1 (DNMT1) is coincident with, but not essential for, global hypomethylation in human placenta. J Biol Chem. 2010 Mar 26;285(13):9583-93. DOI: 10.1074/jbc.M109.064956.

111. Bianco-Miotto T, Mayne BT, Buckberry S, Breen J, Lopez CMR, Roberts CT. Recent progress towards understanding the role of DNA methylation in human placental development. Reproduction (Cambridge, England). 2016;152(1):R23-R30.

112. Vlahos A, Mansell T, Saffery R, Novakovic B. Human placental methylome in the interplay of adverse placental health, environmental exposure, and pregnancy outcome. PLoS genetics. 2019;15(8):e1008236.

113. Price EM, Cotton AM, Peñaherrera MS, McFadden DE, Kobor MS, Robinson W. Different measures of “genome-wide” DNA methylation exhibit unique properties in placental and somatic tissues. Epigenetics. 2012;7(6):652-63.

114. Yuen RK, Penaherrera MS, Von Dadelszen P, McFadden DE, Robinson WP. DNA methylation profiling of human placentas reveals promoter

hypomethylation of multiple genes in early-onset preeclampsia. European Journal of Human Genetics. 2010;18(9):1006.

115. Chu T, Bunce K, Shaw P, Shridhar V, Althouse A, Hubel C, et al. Comprehensive analysis of preeclampsia-associated DNA methylation in the placenta. PLoS One. 2014;9(9):e107318.

116. Leavey K, Bainbridge SA, Cox BJ. Large scale aggregate microarray analysis reveals three distinct molecular subclasses of human preeclampsia. PLoS One. 2015;10(2):e0116508. DOI: 10.1371/journal.pone.0116508.

117. Yeung KR, Chiu CL, Pidsley R, Makris A, Hennessy A, Lind JM. DNA methylation profiles in preeclampsia and healthy control placentas. Am J Physiol Heart Circ Physiol. 2016 May 15;310(10):H1295-303. DOI: 10.1152/ajpheart.00958.2015.

118. Roifman M, Choufani S, Turinsky AL, Drewlo S, Keating S, Brudno M, et al. Genome-wide placental DNA methylation analysis of severely growth-discordant monochorionic twins reveals novel epigenetic targets for intrauterine growth restriction. Clinical epigenetics. 2016;8(1):70.

119. Ruchat S-M, Houde A-A, Voisin G, St-Pierre J, Perron P, Baillargeon J-P, et al. Gestational diabetes mellitus epigenetically affects genes predominantly involved in metabolic diseases. Epigenetics. 2013;8(9):935-43.

120. Finer S, Mathews C, Lowe R, Smart M, Hillman S, Foo L, et al. Maternal gestational diabetes is associated with genome-wide DNA methylation variation in placenta and cord blood of exposed offspring. Human molecular genetics. 2015;24(11):3021-9.

121. Frost JM, Moore GE. The importance of imprinting in the human placenta. PLoS genetics. 2010;6(7):e1001015.

122. Ishida M, Moore GE. The role of imprinted genes in humans. Molecular aspects of medicine. 2013;34(4):826-40.

123. Fowden A, Coan P, Angiolini E, Burton G, Constancia M. Imprinted genes and the epigenetic regulation of placental phenotype. Progress in biophysics and molecular biology. 2011;106(1):281-8.

124. Monk D. Genomic imprinting in the human placenta. American journal of obstetrics and gynecology. 2015;213(4):S152-S62.

125. Pilvar D, Reiman M, Pilvar A, Laan M. Parent-of-origin-specific allelic expression in the human placenta is limited to established imprinted loci and it is stably maintained across pregnancy. Clinical Epigenetics. 2019;11(1):94.

126. El Khattabi LA, Backer S, Pinard A, Dieudonné M-N, Tsatsaris V, Vaiman D, et al. A genome-wide search for new imprinted genes in the human placenta identifies DSCAM as the first imprinted gene on chromosome 21. European Journal of Human Genetics. 2019;27(1):49.

127. Lin X, Su J, Chen K, Rodriguez B, Li W. Sparse conserved under-methylated CpGs are associated with high-order chromatin structure. Genome biology. 2017;18(1):163.

128. Fortin J-P, Triche Jr TJ, Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2016;33(4):558-60.

129. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics. 2016;8(3):389-99.

130. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome biology. 2016;17(1):208.

131. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011 Oct;98(4):288-95. DOI: 10.1016/j.ygeno.2011.07.007.

132. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011 10//;98(4):288-95. DOI: https://doi.org/10.1016/j.ygeno.2011.07.007.

133. Wright ML, Dozmorov MG, Wolen AR, Jackson-Cook C, Starkweather AR, Lyon DE, et al. Establishing an analytic pipeline for genome-wide DNA methylation. Clinical epigenetics. 2016;8(1):45.

134. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-9.

135. Müller F, Scherer M, Assenov Y, Lutsik P, Walter J, Lengauer T, et al. RnBeads 2.0: comprehensive analysis of DNA methylation data. Genome biology. 2019;20(1):55.

136. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics. 2017;33(24):3982-4.

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

138. Marabita F, Almgren M, Lindholm ME, Ruhrmann S, Fagerström-Billai F, Jagodic M, et al. An evaluation of analysis pipelines for DNA methylation profiling using the Illumina HumanMethylation450 BeadChip platform. Epigenetics. 2013;8(3):333-46.

139. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC bioinformatics. 2010;11(1):587.

140. Wahl S, Fenske N, Zeilinger S, Suhre K, Gieger C, Waldenberger M, et al. On the potential of models for location and scale for genome-wide DNA methylation data. BMC bioinformatics. 2014;15(1):232.

141. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2005;54(3):507-54.

142. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research. 2015;43(7):e47-e.

143. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics. 2016;32(2):286-8. DOI: 10.1093/bioinformatics/btv560.

144. Gagnon-Bartsch JA, Jacob L, Speed TP. Removing unwanted variation from high dimensional data with negative controls. Berkeley: Tech Reports from Dep Stat Univ California. 2013:1-112.

145. Weinhold L, Wahl S, Pechlivanis S, Hoffmann P, Schmid M. A statistical model for the analysis of beta values in DNA methylation studies. BMC bioinformatics. 2016;17(1):480.

146. Mansell G, Gorrie-Stone TJ, Bao Y, Kumari M, Schalkwyk LS, Mill J, et al. Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. BMC genomics. 2019;20(1):366.

147. Chen D-P, Lin Y-C, Fann CS. Methods for identifying differentially methylated regions for sequence-and array-based data. Briefings in functional genomics. 2016;15(6):485-90.

148. Jaffe AE, Murakami P, Lee H, Leek JT, Fallin MD, Feinberg AP, et al. Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. International journal of epidemiology. 2012;41(1):200-9.

149. Shaffer JP. Multiple hypothesis testing. Annual review of psychology. 1995;46(1):561-84.

150. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, Lord RV, et al. De novo identification of differentially methylated regions in the human genome. Epigenetics & chromatin. 2015;8(1):6.

151. Butcher LM, Beck S. Probe Lasso: a novel method to rope in differentially methylated regions with 450K DNA methylation data. Methods. 2015;72:21-8.

152. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet. 2006 Dec;38(12):1378-85. DOI: 10.1038/ng1909.

153. Satterthwaite FE. An approximate distribution of estimates of variance components. Biometrics bulletin. 1946;2(6):110-4.

154. Lent S, Xu H, Wang L, Wang Z, Sarnowski C, Hivert M-F, et al. Comparison of novel and existing methods for detecting differentially methylated regions. BMC genetics. 2018;19(1):84.

155. Mallik S, Odom GJ, Gao Z, Gomez L, Chen X, Wang L. An evaluation of supervised methods for identifying differentially methylated regions in Illumina methylation arrays. Briefings in bioinformatics. 2019;20(6):2224-35.

156. Plongthongkum N, Diep DH, Zhang K. Advances in the profiling of DNA modifications: cytosine methylation and beyond. Nature Reviews Genetics. 2014;15(10):647.

157. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. nature. 2009;462(7271):315.

158. Burger L, Gaidatzis D, Schübeler D, Stadler MB. Identification of active regulatory regions from DNA methylation data. Nucleic acids research. 2013;41(16):e155-e.

159. Schroeder DI, Lott P, Korf I, LaSalle JM. Large-scale methylation domains mark a functional subset of neuronally expressed genes. Genome research. 2011;21(10):1583-91.

160. Kishore K, de Pretis S, Lister R, Morelli MJ, Bianchi V, Amati B, et al. methylPipe and compEpiTools: a suite of R packages for the integrative analysis of epigenomics data. BMC bioinformatics. 2015;16(1):313.

161. Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, et al. Human body epigenome maps reveal noncanonical DNA methylation variation. Nature. 2015;523(7559):212.

162. Frost JM, Moore GE. The importance of imprinting in the human placenta. PLoS genetics. 2010;6(7).

163. Hanna CW, Peñaherrera MS, Saadeh H, Andrews S, McFadden DE, Kelsey G, et al. Pervasive polymorphic imprinted methylation in the human placenta. Genome research. 2016;26(6):756-67.

164. Rezwan FI, Docherty LE, Poole RL, Lockett GA, Arshad SH, Holloway JW, et al. A statistical method for single sample analysis of HumanMethylation450 array data: genome-wide methylation analysis of patients with imprinting disorders. Clinical epigenetics. 2015;7(1):48.

165. Santoni FA, Stamoulis G, Garieri M, Falconnet E, Ribaux P, Borel C, et al. Detection of imprinted genes by single-cell allele-specific gene expression. The American Journal of Human Genetics. 2017;100(3):444-53.

166. Green BB, Kappil M, Lambertini L, Armstrong DA, Guerin DJ, Sharp AJ, et al. Expression of imprinted genes in placenta is associated with infant neurobehavioral development. Epigenetics. 2015;10(9):834-41.

167. Heiss JA, Just AC. Improved filtering of DNA methylation microarray data by detection p values and its impact on downstream analyses. Clinical epigenetics. 2019;11(1):15.

168. Heiss JA, Just AC. Identifying mislabeled and contaminated DNA methylation microarray data: an extended quality control toolset with examples from GEO. Clinical epigenetics. 2018;10(1):73.

169. Wan Q, Leemaqz SY-L, Pederson SM, McCullough D, McAninch DC, Jankovic-Karasoulos T, et al. Quality control measures for placental sample purity in DNA methylation array analyses. Placenta. 2019;88:8-11.

170. Strimbu K, Tavel JA. What are biomarkers? Current Opinion in HIV and AIDS. 2010;5(6):463.

171. Manokhina I, Del Gobbo GF, Konwar C, Wilson SL, Robinson WP. Placental biomarkers for assessing fetal health. Human molecular genetics. 2017;26(R2):R237-R45.

172. Wilson SL, Robinson WP. Utility of DNA methylation to assess placental health. Placenta. 2018;64:S23-S8.

173. Leavey K, Wilson SL, Bainbridge SA, Robinson WP, Cox BJ. Epigenetic regulation of placental gene expression in transcriptional subtypes of preeclampsia. Clinical epigenetics. 2018;10(1):28.

174. Van Schendel RV, Kleinveld JH, Dondorp WJ, Pajkrt E, Timmermans DR, Holtkamp KC, et al. Attitudes of pregnant women and male partners towards non-invasive prenatal testing and widening the scope of prenatal screening. European Journal of Human Genetics. 2014;22(12):1345.

175. Wu P, Van den Berg C, Alfirevic Z, O'Brien S, Röthlisberger M, Baker PN, et al. Early pregnancy biomarkers in pre-eclampsia: a systematic review and meta-analysis. International journal of molecular sciences. 2015;16(9):23035-56.

176. Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. Proceedings of the National Academy of Sciences. 2008;105(42):16266-71.

177. Godfrey KM, Costello PM, Lillycrop KA. The developmental environment, epigenetic biomarkers and long-term health. Journal of developmental origins of health and disease. 2015;6(5):399-406.

178. Leenen FA, Muller CP, Turner JD. DNA methylation: conducting the orchestra from exposure to phenotype? Clinical epigenetics. 2016;8(1):92.

179. Horvath S. DNA methylation age of human tissues and cell types. Genome biology. 2013;14(10):3156.

180. Dhingra R, Kwee LC, Diaz-Sanchez D, Devlin RB, Cascio W, Hauser ER, et al. Evaluating DNA methylation age on the Illumina MethylationEPIC Bead Chip. PloS one. 2019;14(4):e0207834.

181. Knight AK, Craig JM, Theda C, Bækvad-Hansen M, Bybjerg-Grauholm J, Hansen CS, et al. An epigenetic clock for gestational age at birth based on blood methylation data. Genome biology. 2016;17(1):206.

182. Bohlin J, Håberg SE, Magnus P, Reese SE, Gjessing HK, Magnus MC, et al. Prediction of gestational age based on genome-wide differentially methylated regions. Genome biology. 2016;17(1):207.

183. Lee Y, Choufani S, Weksberg R, Wilson SL, Yuan V, Burt A, et al. Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels. Aging (Albany NY). 2019;11(12):4238.

184. Cotton AM, Lam L, Affleck JG, Wilson IM, Penaherrera MS, McFadden DE, et al. Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation. Human genetics. 2011;130(2):187-201.

185. Jung C-H, Park D, Georgeson P, Mahmood K, Milne R, Southey M, et al. sEst: Accurate Sex-Estimation and Abnormality Detection in Methylation Microarray Data. International journal of molecular sciences. 2018;19(10):3172.

186. Liu J, Morgan M, Hutchison K, Calhoun VD. A study of the influence of sex on genome wide methylation. PloS one. 2010;5(4):e10028.

187. Yuan V, Price EM, Del Gobbo GF, Mostafavi S, Cox B, Binder AM, et al. Accurate ethnicity prediction from placental DNA methylation data. bioRxiv. 2019:618470.

188. Rider CF, Carlsten C. Air pollution and DNA methylation: effects of exposure in humans. Clinical epigenetics. 2019;11(1):131.

189. Ivorra C, Fraga MF, Bayón GF, Fernández AF, Garcia-Vicent C, Chaves FJ, et al. DNA methylation patterns in newborns exposed to tobacco in utero. Journal of translational medicine. 2015;13(1):25.

190. Zhong H, Kim S, Zhi D, Cui X. Predicting gene expression using DNA methylation in three human populations. PeerJ. 2019;7:e6757.

2 Comparison of methods for region-based differential methylation analysis for Illumina EPIC arrays

Abstract

DNA methylation arrays, such as the Illumina Infinium® MethylationEPIC BeadChip, are a cost-effective way to assess DNA methylation across large populations because they cover all the important regulatory elements. A number of Bioinformatics methods have been developed to analyse these arrays. However, there are few studies that comprehensively benchmark the performance of common statistical methods that are used for EPIC arrays. To help researchers choose optimal methods when doing differential methylation analysis, we conducted an evaluation of three methods for identifying differentially methylated regions (DMRs) with particular attention to the investigation of suitable cutoff parameters for defining DMRs and the performance of identification methods. Results showed that DMRcate detected more DMRs than bumphunter and Probe Lasso and the identified DMRs overlapped with DMRs identified by the other two methods. The output results of these methods varied when parameters were changed, especially for Probe Lasso because it uses flexible windows to identify DMRs instead of loess or kernel smoothing methods. Overall, the DMRcate method consistently detected more true DMRs than the other two methods making it a good choice of DMR finder in scenarios where the DNA methylation changes between groups are relatively small.

Keywords: DNA methylation, EPIC array, DMR identification, Comparison

2.1 Introduction

DNA methylation is one of the most studied epigenetic modifications that are heritable traits without the changing of underlying DNA sequences. The most widely characterised DNA methylation process is the addition of a methyl group at the 5th carbon of the cytosine resulting in the 5-methylcytosine (5-mC). When located at gene promoters and enhancers, DNA methylation potentially represses gene expression [1]. DNA methylation also plays important roles in determining tissue-specific gene expression, chromosomal stability, gene imprinting, as well as X chromosome inactivation in females [2]. In human disease, the DNA methylation patterns are altered, especially for cancer cells [3]. Robust and reliable techniques are critical for obtaining the DNA methylation profiles of different cells and tissue types and to make sure these profiles are compatible and allow data mining and result comparison for researchers all around the world [4].

While whole-genome bisulfite sequencing (WGBS) is the gold standard for measuring DNA methylation, the high cost of WGBS limits its use in studies with large sample size. Array-based technologies provide an economical solution to assessing large numbers of samples, being high-throughput and widely adopted [5]. For example, platforms such as the Illumina Infinium® HumanMethylation450 BeadChip (450K array) and the Illumina Infinium® MethylationEPIC BeadChip (EPIC array) are commonly used in DNA methylation experiments. EPIC arrays include more than ninety percent of the 450K array probes [4] and have almost twice the number of cytosine positions (865859 sites) with these additional probes targeting sites in regulatory regions identified by ENCODE [6] and FANTOM5 [7] projects.

In a case-control experiment, differences between conditions can be assessed by identifying Differentially Methylated Position (DMP) and Differentially Methylated Regions (DMRs) defining a single CpG site or entire region displaying differences. The identification of DMPs and DMRs may be associated with a change in phenotype and can be either directly or indirectly linked to the variance in gene expression [8]. DMPs were widely used as biomarkers for disease, non-invasive testing or for predicting age [9-11]. Except for DMPs at SNPs, DMPs are not always as biologically meaningful as DMRs because DNA methylation changes at gene promoter regions or imprinting control regions may canonically influence gene expression [12]. Many DMRs have been found in the developmental stages involved in cell proliferation and differentiation [13] and in the cell reprogramming process [14]. Some DMRs are related to tissue-specific gene expression [8]. Therefore, it is critical and fundamental to detect DMRs, especially those in functional regions such as promoters and enhancers that may be associated with transcriptional regulation.

Methods for identifying DMRs include aggregating the information of DNA methylation sites within a predefined region (such as *Probe Lasso*) or directly defining DMRs with regression models (such as *DMRcate* and *bumphunter*) [15]. In this study, we investigated the performance of methods for detecting DMRs, and identified important advantages and shortcomings of each method. These will provide insights into how to choose an appropriate method for differential methylation analysis using EPIC array data.

2.2 Materials and Methods

2.2.1 Ethics statement

Prior to collection of placental tissue samples, written, informed consent was obtained from all subjects involved in this study. Collection of placental tissue from elective termination of first and second trimester pregnancies was approved by the Queen Elizabeth Hospital and Lyell McEwin Hospital Human Research Ethics Committee (TQEH/LMH/MH and HREC/16/TQEH/33).

2.2.2 Sample collection

Placenta samples were obtained from elective terminations from the Pregnancy Advisory Centre at Woodville, South Australia. The gestational age was determined using transabdominal ultrasonography. Females with infection, endocrine abnormalities, antiphospholipid syndrome or other known complications were excluded from the study. Placenta samples were collected and dissected within 15 minutes of termination of the pregnancies and the regions of villous tissue were isolated [16]. The same type of tissue (placental chorionic villous tissue) was collected from all women bearing in mind the tissue specificity of DNA methylation [17]. Villous tissue for DNA extraction was isolated and then being stored at -80°C until DNA was extracted.

2.2.3 Array processing

DNA was extracted from placental tissue samples using a modified version of the TES protocol [18]. For each sample, 1 µg of DNA was sent to PathWest Laboratory Medicine (QEII Medical Centre, Perth, Western Australia) for bisulfite-conversion and hybridisation to the Illumina Infinium® MethylationEPIC BeadChip according to the manufacturer's instructions.

2.2.4 Algorithms for identifying differentially methylated regions (DMRs)

Three methods were tested for identifying DMRs (Table 2.1). *Bumphunter*, *Probe Lasso* and *DMRcate* methods are all based on linear models, so M values (logit transformed percentage of DNA methylation) instead of beta values (percentage of DNA methylation) are chosen as input values. The difference of these three methods is focusing on how the aggregate significant probes into regions. In *Bumphunter* method, probes were separated into clusters based on their positions on genome, then the M values of each probe in their specific position was fitted in a linear model ($Y_{ij} = \mu(t_j) + \beta(t_j)X_i + \gamma(t_j)Z_i + \alpha_j W_i + \varepsilon_{i_j}$)[1] to estimate $\beta(t_j)$. $\beta(t_j)$ was then used to estimate the loess smooth function $\beta(t)$ with each point weighted. Candidate regions were generated using contiguous runs of measurements for which $\beta(t) > K$ or $\beta(t) < K$ where K could be a user defined threshold. Statistical uncertainty of each estimated

[1] The variable t_j denotes the location on the genome of the j-th locus, i denotes each individual probe, Y_{ij} is the M value of one probe at j-th location. $\mu(t_j)$ is the average DNA methylation of control group, X is the variable of interest, Z is the known confounder, W is the unmeasured confounder estimated by *sva* package, and the residuals are assumed to be normally distributed with $\varepsilon_{ij} \sim N(0, \sigma^2(t_j))$.

DMR were tested using bootstrap or permutation tests as the last step [19]. Family wise error rate (FWER) [20] of each DMR represented the significance of DMR (by comparing permutated DMRs to the observed DMRs).

Probe Lasso method first uses M values to fit into a linear model for differentially methylated sites. As the second step, flexible windows for identifying DMRs were generated according to genetic and epigenetic annotations of each probe on the EPIC array. Dynamic window/lasso sizes were assigned to each category based on the difference of probe spacing features between these categories [21]. Overlapping and neighbouring of lassos (or windows) were merged if they were separated by less than the user specified threshold. A DMR was identified when merging ceased. Stouffer's method was used to estimate p-values for DMRs.

In *DMRcate* method, standard linear models were first applied to M values to analyse differential methylation at probe level. The signed t statistics between two groups from *limma* for each site were transferred to unsigned F statistics (t^2). The Gaussian smoothed test statistics were then modelled by a scaled chi-squared random variable $a_i\chi^2_{b_i}$ (a_i and b_i were constants chose to match the mean and variance of smoothed test statistics). For each site, two smoothed estimates were computed: one weighted and one not, for a null comparison. The two estimates were compared via a Satterthwaite approximation [22]. Finally, nearby significant sites that were at most λ bp (λ is a use defined value) from each other and had FDR-corrected p-values smaller than 0.05 were agglomerated as DMRs [23]. Stouffer's method was used to estimate p-value of each DMR.

Table 2.1 Methods for identifying DMRs.

Name	Package	Package URL	Program language	Ref
Bumphunter	*minfi*	https://bioconductor.org/packages/release/bioc/html/minfi.html	R	[19]
Probe Lasso	*ChAMP*	https://bioconductor.org/packages/release/bioc/html/ChAMP.html	R	[21]
DMRcate	*DMRcate*	https://bioconductor.org/packages/release/bioc/html/DMRcate.html	R	[23]

2.2.5 Visualisation of results

The visualisation of the results in this study was accomplished using R packages *VennDiagram* [24] and *ggplot2* [25].

2.2.6 Analyses code available:

Data was analysed on a 2016 MacBook Pro running a 2.9 GHz Intel Core i5 processor with 8 GB 2133 MHz LPDDR3. Data were processed using programming language R with R version 3.6.0 (2019-04-26) [26]. The codes related with this study are available in GitHub: https://github.com/QianhuiWan/Methods_DNAm.

2.3 Results

2.3.1 Selection of samples and parameters used in each method

The public data sets including control and *NNMT* overexpressed samples of WI-38 fibroblast cell line were obtained from GEO database (GSE126672) [27] because the samples from a cell line were considered to have more similarities between replicates

than placenta samples from different individuals. The performance of DMR methods were first tested based on this dataset because there is a large DNA methylation change between the control and overexpressed groups. All data were from the Illumina Infinium® MethylationEPIC platform (Table 2.2). EPIC array data of twelve placenta samples from first and second trimester (Table 2.3) were selected and used to verify the result obtained from public data sets.

We tested a variety of functions and parameters listed in Table 2.4 for identifying DMRs. First, DMRs were detected using default parameters of each method. Second, according to a previous study [28], additional optimised parameters (underlined in Table 2.4) of each method were used for DMR analyses. Comparable parameters with the setting of 500 bp as the minimum distance (bp) between neighbouring DMRs were also compared using these data. Third, the comparable parameters with the setting of 500 bp as the minimum distance (bp) between neighbouring DMRs were also tested to compare these methods more closely.

Table 2.2 Public data sets used in this study.

GEO accession	Cell line
GSM3611046	WI-38 fibroblasts control
GSM3611047	WI-38 fibroblasts control
GSM3611048	WI-38 fibroblasts overexpressing *NNMT*
GSM3611049	WI-38 fibroblasts overexpressing *NNMT*

Table 2.3 Placenta samples used for this study.

Tissue	Batch	Trimester	Fetal sex	Gestation (weeks')
Placenta	1	First	F	8
Placenta	1	Second	F	14
Placenta	1	Second	M	16
Placenta	1	Second	M	16
Placenta	1	Second	M	13

Placenta	1	First	F	6
Placenta	1	First	F	7
Placenta	2	Second	F	15
Placenta	2	Second	F	14
Placenta	2	First	M	8
Placenta	2	First	F	6
Placenta	2	First	M	7

Table 2.4 Functions and parameters used for comparing methods.

Name	Function	Default parameters	Optimised parameters	Comparable parameters
Bumphunter	bumphunter()	nullMethod="bootstrap", smooth=FALSE, useWeights=FALSE, maxGap=500, pickCutoffQ=0.99, B=5, minNum=7	nullMethod="bootstrap", smooth= TRUE, useWeights= TRUE, maxGap=250, pickCutoffQ=0.95, B=5, minNum=2	nullMethod="bootstrap", smooth=TRUE, useWeights=TRUE, maxGap=500, pickCutoffQ=0.99, B=5, minNum=2
Probe Lasso	Champ.DMR()	minDmrSep=1000, meanLassoRadius=375, minProbes=7	minDmrSep=1000, meanLassoRadius=1000, minProbes=2	minDmrSep=500, meanLassoRadius=1000, minProbes=2
DMRcate	dmrcate()	lambda = 1000, C = 2, min.cpgs = 2	lambda = 500, C = 5, min.cpgs = 2	lambda = 500, C = 5, min.cpgs = 2

2.3.2 EPIC array data analysis

2.3.2.1 Quality control

The quality of the EPIC array data was assessed using *minfi* package. Scatterplots of median unmethylation (Unmeth) signal versus median methylation (Meth) signal values were generated for assessing quality for each sample. Good samples clustered together, while failed samples tend to separate and had lower median intensities [29]. In addition, the sample-dependent control probes in the EPIC array were used to check the quality of bisulfite conversion for type I and II assays, non-polymorphic performance, specificity of matching type I and II probes and the system background.

2.3.2.2 Probe filtering and normalisation

Failed and unwanted probes were filtered for data from the WI-38 cell line. The removed failed probes included 1466 probes with detection P > 0.01, 2806 probes with probes < 3 beads in 5% of the 4 samples, 167592 probes with SNPs at CpG/SBE sites on probes [29], and 42999 cross-reactive probes [30]. Probes on sex chromosomes were kept for samples from WI-38 cell line because they are all female. In total, 214863 probes were removed, and 651973 probes remained for subsequent processing.

For 12 placenta samples the failed probes removed included 4877 probes with detection P > 0.01, 40128 probes with probes < 3 beads in 5% of the 12 samples, 157805 probes with SNPs at CpG/SBE/SNP sites on probes [29], and 41365 cross-reactive probes [30]. 15181 probes on X/Y chromosomes were also removed to decrease the disturbance variables. In total, 259356 probes were removed, and 606530 probes remained for subsequent processing.

Filtered data were pre-processed with *ENmix* package to remove background noise and correct dye bias [31]. Regression on logarithm of internal control probes (RELIC) from *Enmix* package was used for dye bias correction to improve accuracy of methylation beta value estimates [32]. The corrected data were normalised with Beta-mixture quantile (BMIQ) normalisation method implemented in *ChAMP* package [33]. The normalised data were used for downstream analysis.

2.3.2.3 Method comparison workflow

Pre-processing and DMR analyses of EPIC array data were accomplished and Figure 2.1 shows the overall workflow of the analyses in this study. After data pre-processing and normalisation, the three methods were used for identifying DMRs. Different aspects of these methods were compared, including number of true DMRs identified, number of DMRs overlapped with promoters and enhancers, length of the detected DMRs, common DMRs identified across different methods and time used for identifying DMRs.

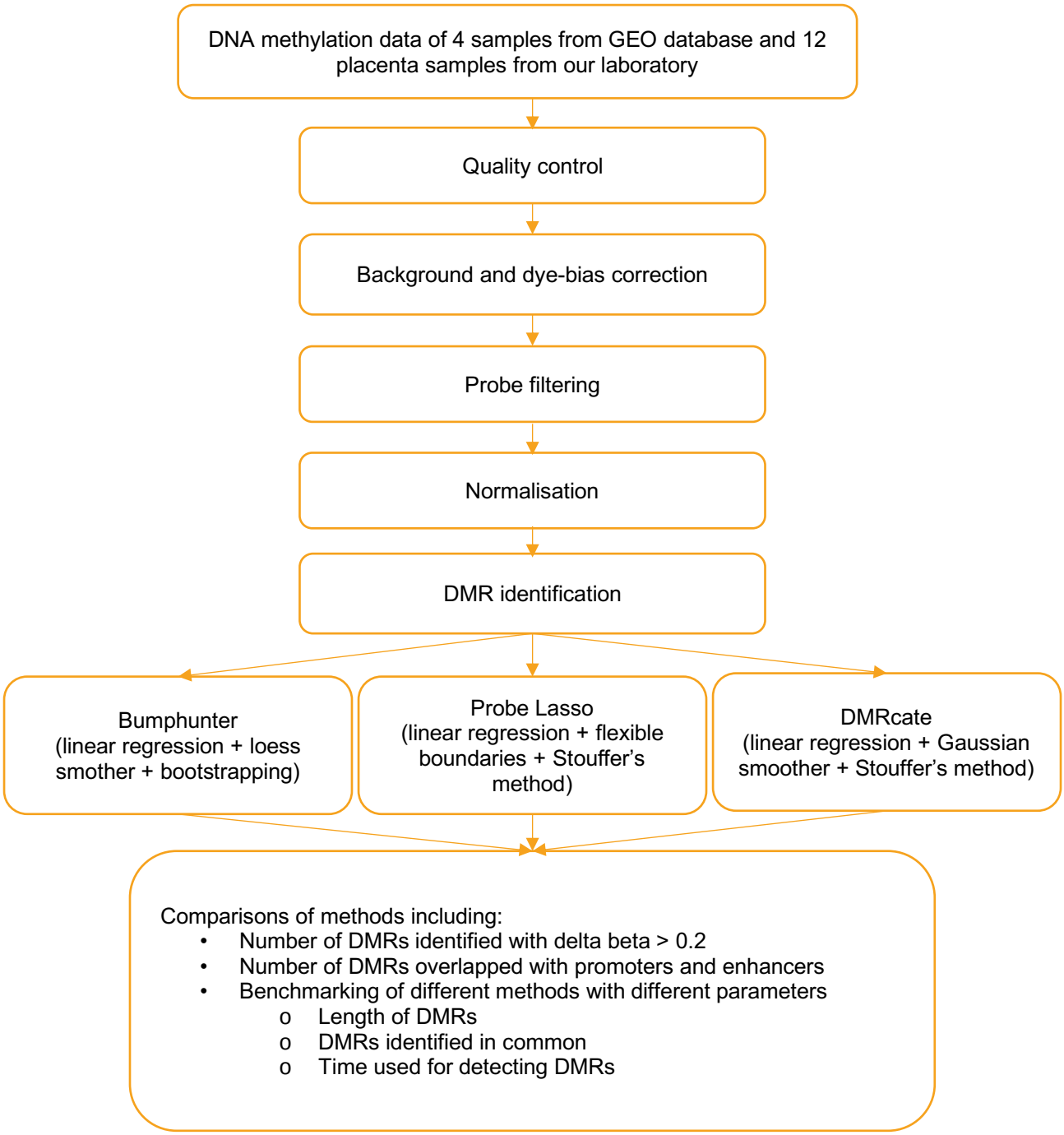

Figure 2.1 Overall workflow of the analyses.

2.3.3 Accuracy and annotation of identified DMRs

Since change of mean beta values (percentage of DNA methylation) higher than 0.2 ($|\text{mean}\Delta\beta| > 0.2$) between groups were considered as biologically meaningful DNA

methylation changes detected by EPIC arrays [34, 35], the number of DMRs (FWER or Stouffer's p values < 0.05) identified with mean beta value differences greater than 0.2 were considered as true DMRs in this study. From the detailed workflow using data from WI-38 cells, *bumphunter* with default parameters had 99 DMRs called, which decreased to 5 and 3 respectively when optimised and comparable parameters were used (Table 2.4). *Probe Lasso* method called 3 DMRs when using default parameters, which increased to 3180 and 3211 when optimised and comparable parameters were used (Table 2.4). *DMRcate* detected 3529 DMRs with default parameters and the number increased to 4516 when using optimised and comparable parameters (Table 2.4).

All DMRs (100%) called by *bumphunter* with default and comparable parameters have mean beta values changes higher than 0.2, which decreased to 40% when using optimised parameters. Probe Lasso method has 2 out of 3 DMRs (67%) with mean beta values changes higher than 0.2 when using default parameters. This decreased to 4.2% and 4.1% when optimised and comparable parameters were used. DMRcate detected 395 out of 3529 DMRs (11%) with mean beta values changes higher than 0.2 when using default parameters and the percentage increased to 15% when using optimised and comparable parameters.

In order to know whether these true DMRs ($|\text{mean}\Delta\beta| > 0.2$) overlapped with regulatory regions such as promoters and enhancers, the number of DMRs overlapped with human promoters and enhancers was calculated. Bumphunter with default parameters had 9 DMRs overlapped with promoters, which decreased to 2 when optimised and comparable parameters were used. There was no DMR identified by

Bumphunter that overlapped with enhancers. Probe Lasso method has 1 DMR that overlapped with promoters when using default parameters, which increased to 35 when optimised and comparable parameters were used. There are 7 DMRs identified by Probe Lasso that overlapped with enhancers. DMRcate detected 34 DMRs that overlapped with promoters and 4 DMRs overlapped with enhancers with default parameters. The number of DMRs that overlapped with promoters increased to 41 and the number of DMRs overlapped with enhancers decreased to 1 when using optimised and comparable parameters (Table 2.5).

The results show that Probe Lasso detected less percentage of DMRs with mean beta value difference greater than 0.2 than DMRcate and bumphunter. Bumphunter identified fewer DMRs but the percentage of true DMRs is higher than for the other two methods. The results show that there were no DMRs that overlapped with enhancers when using bumphunter method. More DMRs on promoters rather than enhancers were detected by these methods.

Table 2.5 Comparison of accuracy and annotation of identified DMRs.

Name	Statistic used for calculating p values for DMRs	Parameters	Number of DMRs	Number of DMRs (|mean$\Delta\beta$ | > 0.2)	Number of DMRs overlapped with promoters	Number of DMRs overlapped with enhancers
bumphunter	Familywise error rate (FWER)	Default	99	99 (100%)	9	0
		Optimised	5	2 (40%)	2	0
		Comparable	3	3 (100%)	2	0
Probe Lasso	Stouffer's method	Default	3	2 (67%)	1	0
		Optimised	3180	133 (4.2%)	35	7
		Comparable	3211	133 (4.1%)	35	7
DMRcate	Stouffer's method	Default	3529	395 (11%)	34	4
		Optimised	4516	691 (15%)	41	1
		Comparable	4516	691 (15%)	41	1

2.3.4 Length of DMRs identified by different methods

The variation of the length of identified DMRs using different methods was compared as one aspect reflecting the accuracy of methods. DMR length identified by bumphunter is shorter when using the default parameter than using optimised and comparable parameters. DMR length identified by Probe Lasso is more variable with optimised and comparable parameters than default parameters. DMR length identified by DMRcate is stable between three different parameter settings. DMRs identified by Probe Lasso are always longer than DMRs identified by the other two methods and the changing of parameters influenced Probe Lasso more than Bumphunter and DMRcate (Figure 2.2).

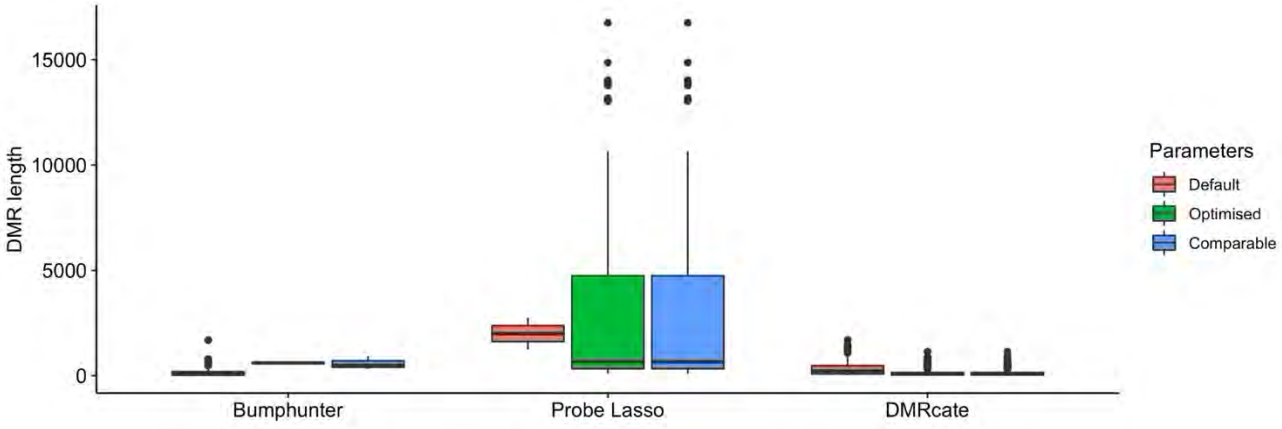

Figure 2.2 Length of DMRs identified by the three methods under default, optimised and comparable parameters.

2.3.5 Overlapped regions identified by the three methods

DMRs identified by different methods were not entirely the same but had overlapping regions. The overlapping of DMRs detected by three methods with P value smaller than 0.05 and mean beta value changes greater than 0.2 were investigated. When

using default parameters, there were 39 overlapped regions between DMRs identified by Bumphunter and DMRcate, 2 DMRs identified by Probe Lasso all overlapped with DMRs identified by DMRcate and there were no DMRs identified in common by all three methods (Figure 2.3A). The proportion of overlapped DMRs increased when using the optimised and comparable parameters, 74 DMRs were identified by both Probe Lasso and DMRcate, but only 1 DMR was identified by all three methods (Figure 2.3B-C).

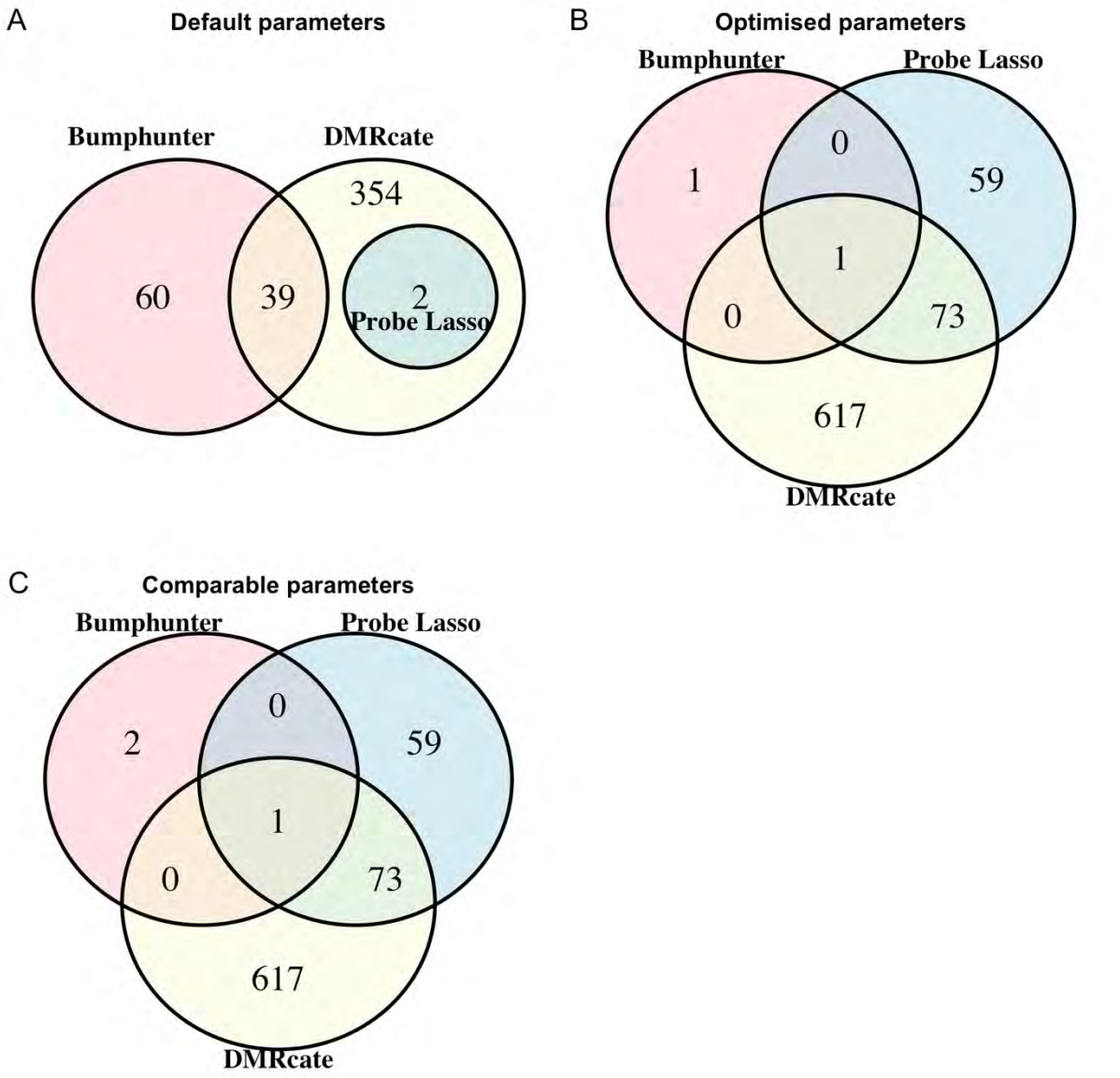

Figure 2.3 Overlap in identified DMRs using the three methods with default (A), optimised (B) and comparable (C) parameters.

2.3.6 Benchmarking of each method for DMR identification

To investigate which method for identifying DMRs is more computationally efficient, we tested the time used for each method using the public data sets of the cell line (Table 2.5 and Figure 2.4). All 651,973 EPIC array sites after the filtering of failed probes were used for benchmarking and the R package *microbenchmark* [36] calculated the mean time used for each method by repeating each method 5 times. Under default parameters, the mean time used to detect DMRs is 0.05, 0.0043 and 0.012 hours, respectively for bumphunter, Probe Lasso and DMRcate. The time increased to 0.64, 0.0079 and 0.013 hours when using optimised parameters, and slightly increased to 0.70, 0.0080 hours for bumphunter and Probe Lasso when comparable parameters were used (Table 2.6). Bumphunter took more time for processing than Probe Lasso and DMRcate which could be because of the bootstrapping step for estimating significance of each DMR [19]; the time for analysis using bumphunter varies the most with the changing of parameters (Figure 2.4).

Table 2.6 Time used for identifying DMRs with different methods.

Name	Parameters	Mean time (hours)	Size of the output value
bumphunter	Default	0.05	68.7 MB
	Optimised	0.64	63 MB
	Comparable	0.70	63.2 MB
Probe Lasso	Default	0.0043	6.7 KB
	Optimised	0.0079	951 KB
	Comparable	0.0080	959 KB
DMRcate	Default	0.012	845 KB
	Optimised	0.013	951 KB
	Comparable	0.013	951 KB

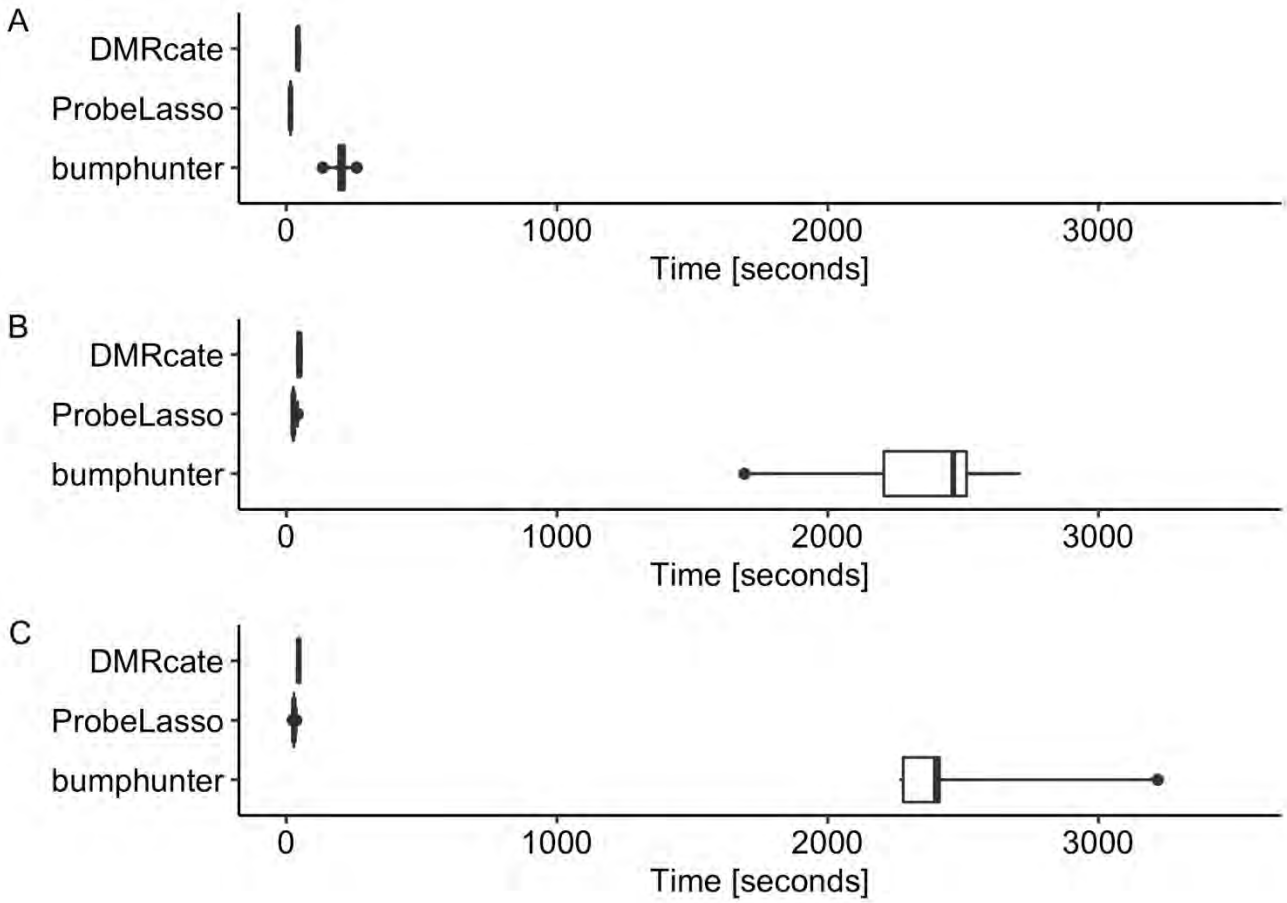

Figure 2.4 Benchmarking of each method for detecting DMRs. The time of using default (A) parameters for each method is shorter than optimised (B) parameters and comparable parameters (C).

2.3.7 Verify the performance of DMR methods using data from placenta samples

By applying the same three sets of parameters to detect DMRs between placenta samples from first and second trimester, we can verify the DMR method comparison results obtained from the public data. When using default parameters, there are 21 overlapped regions between DMRs identified by Bumphunter and DMRcate, no DMRs identified by Probe Lasso and there is no DMR identified in common by all three methods (Figure 2.5A). The proportion of overlapped DMRs increased when using the optimised parameters, 67 DMRs were identified by both Probe Lasso and DMRcate,

no DMRs were identified by Bumphunter and there was no DMR identified in common by all three methods (Figure 2.5B). With comparable parameters, 67 DMRs were identified by Probe Lasso and DMRcate, 2 DMRs identified by Bumphunter were not identified by the other two methods and there was no DMR identified in common by all three methods (Figure 2.5C). DMRcate still identified the highest number of DMRs compared to the other two methods. There are overlapped regions between DMRs identified by different methods and the overlapping pattern is similar to the public data sets. That is, very few DMRs were identified in common by these 3 methods. When using default parameters, Probe Lasso did not identify any DMRs which could be due to the default setting that defines a DMR with at least 7 probes while the optimised and comparable parameters use a minimum of 2 probes cutoff. The number of DMRs identified by bumphunter decreased when changing from default parameters to optimised and comparable parameters because a smoother is used to combine nearby regions when using optimised and comparable parameters.

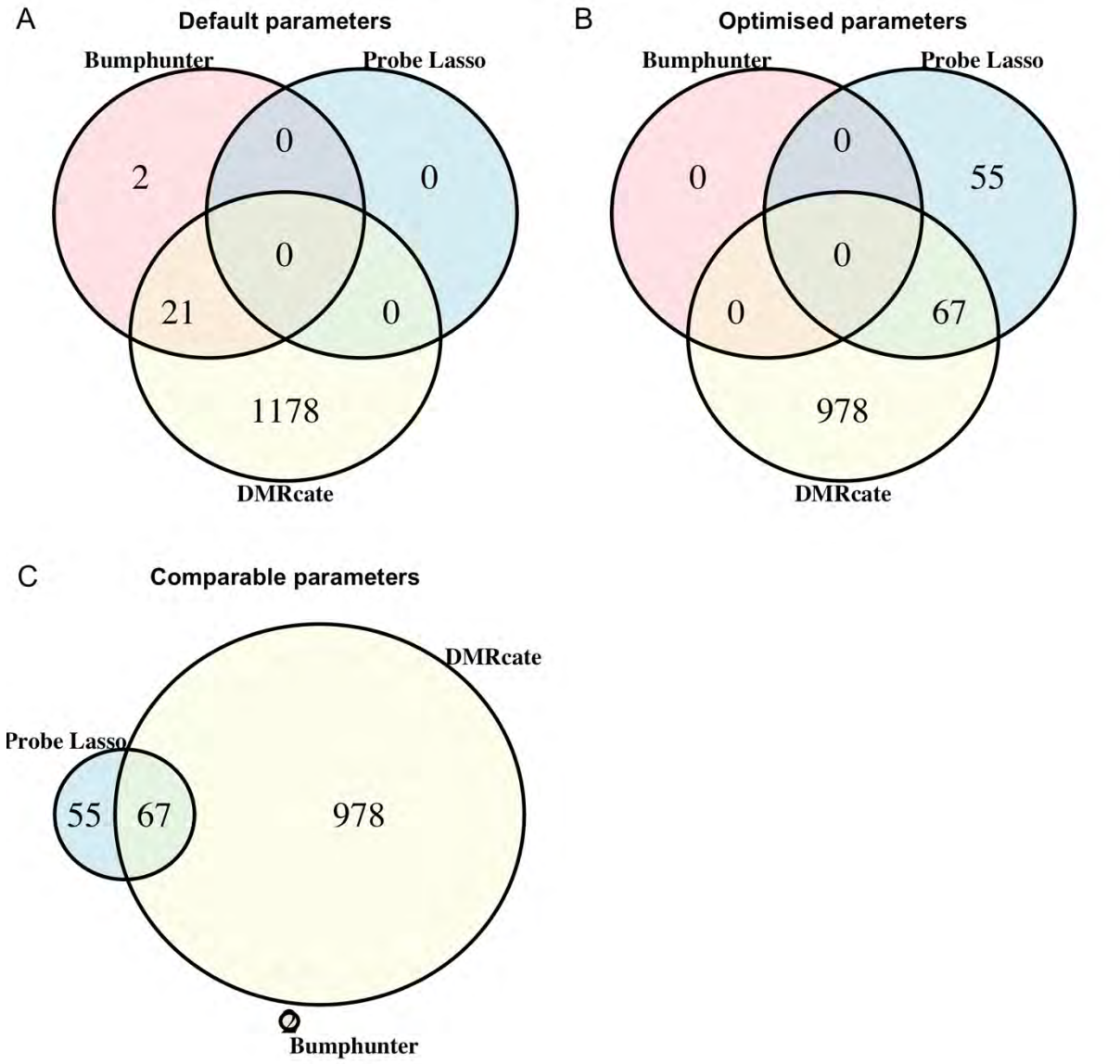

Figure 2.5 Overlapping of DMRs of placenta samples (first trimester vs second trimester) identified by the three methods using default (A), optimised (B) and comparable (C) parameters.

2.4 Discussion

As genetic variation can only explain part of the disease mechanisms, DNA methylation, as one type of epigenetic variation, is promising in detecting regulatory mechanisms associated with environmental exposures and disease risks. So far, the development of genome wide DNA methylation profiling techniques enabled the study

of DNA methylation changes across many samples and tissue types. These DNA methylation data need to be properly analysed and carefully interpreted, because tissue types, tissue purity, batch effects, control factors and disease group need to be considered before and during data processing, and functional relevance needs to be demonstrated by further experiments. Region-based differential methylation analysis between different cell and tissue types can be an important step during data processing. To this end, it is necessary to evaluate the different methods for detecting DMRs. Real DNA methylation EPIC data sets were used instead of using the simulated data sets because these data can represent the methylome of cell lines and tissue types better and avoid some of the bias introduced by simulated data [37].

The different methods with different parameters were first tested by using the public data sets from WI-38 fibroblast cell lines and the two groups in this data sets were control wild-type cells and *NNMT* overexpressed cells [27]. Different from cell lines, placenta samples were bulk tissue and from different individuals with different weeks of gestation. Therefore, it is expected that DNA methylation changes between placenta from first and second trimester were not as predictable as that between control cell line and genetically modified cell line.

Three methods including bumphunter, Probe Lasso and DMRcate for detecting DMRs were compared in this study. These three methods rarely identify precisely the same DMRs that have the same start and end bases, since these methods differ in the assumptions made and statistical approaches taken [38]. DMRcate was considered more sensitive to DMRs because it uses unsigned F statistics and Gaussian smoothing [39]. This is more suitable for placenta samples than the other two methods

because the DNA methylation changes in placenta are not as dramatic as changes between control and overexpressed cell lines. Bumphunter can take known confounders and unmeasured confounders as covariates in linear regression, and along with a bootstrap analysis on DMR uncertainty, the program can have a longer processing time. The Probe Lasso method tends to be more biologically meaningful because it has the advantage of using sliding windows based on the biological features such as promoters, UTRs and gene bodies, and uses varying size of the lasso encompassing them [21]. However, Probe Lasso is not able to adjust batch effects and other confounders in the models and it can split some DMRs that straddle multiple annotations which could decrease the sensitivity of the method [23]. All three methods above were all powerful enough to detect changes in methylation patterns in small regions (DMR<2Kb) but they had limited ability to detect large DMRs [39], since large DMRs can have lower mean beta value differences between groups that are hard to detect.

Identification of DMRs is more data dependent and complex than identifying DMPs, as previous studies have detailed, and all three methods are designed for identifying DMRs between categorical variables, but they are not suitable for detecting small and consistent DNA methylation changes across samples (Mallik et al., 2019). There are potential ways that DMR finding may be improved. First, the researchers can filter the sites according to the regions that they are focusing on. For example, they can use the sites in gene promoters, which can reduce false positive DMRs detected in the regions that are not of interest. Second, if the purpose of the study is to identify DNA methylation changes in large regions, *blockFinder* for identifying differentially methylated genomic blocks (DMB) is recommended. Third, using EPIC array with data

from more DNA methylation sites will potentially reduce the false positive DMRs since the estimate of mean beta difference in a region with more sites is more accurate than a region with few sites detected. In addition, there are more changing parameters that need to be considered when identifying DMRs than identifying DMPs, so it is better for researchers to test different parameters and choose the suitable parameters to use in their study.

In summary, more data sets and simulated datasets could be needed for refining the parameters for EPIC array data. The optimised DMR parameters perform best for 450K array probably not all suitable for EPIC arrays since EPIC array has more probes located in the open sea regions. This study will help researchers to select the appropriate methods for detecting DMRs in their studies.

2.5 Conclusion

This study compared methods for detecting DMRs using EPIC array and recommended using *DMRcate* for DMR analyses. For detecting large differentially methylated blocks and small and consistent changes, other methods need to be considered.

References

1. Baribault C, Ehrlich KC, Ponnaluri VC, Pradhan S, Lacey M, Ehrlich M. Developmentally linked human DNA hypermethylation is associated with down-modulation, repression, and upregulation of transcription. Epigenetics. 2018;13(3):275-89.
2. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nature Reviews Genetics. 2010;11(3):204-20.
3. Mansell G, Gorrie-Stone TJ, Bao Y, Kumari M, Schalkwyk LS, Mill J, et al. Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. BMC genomics. 2019;20(1):366.
4. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics. 2016;8(3):389-99.
5. Wang T, Guan W, Lin J, Boutaoui N, Canino G, Luo J, et al. A systematic study of normalization methods for Infinium 450K methylation data using whole-genome bisulfite sequencing data. Epigenetics. 2015;10(7):662-9.
6. Dunham I, Birney E, Lajoie BR, Sanyal A, Dong X, Greven M, et al. An integrated encyclopedia of DNA elements in the human genome. 2012.
7. Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome biology. 2015;16(1):22.
8. Rakyan VK, Down TA, Thorne NP, Flicek P, Kulesha E, Gräf S, et al. An integrated resource for genome-wide identification and analysis of human

tissue-specific differentially methylated regions (tDMRs). Genome research. 2008;18(9):1518-29.

9. Cummings JL. Biomarkers in Alzheimer's disease drug development. Elsevier; 2011.

10. Gallardo-Gómez M, Moran S, de la Cadena MP, Martínez-Zorzano VS, Rodríguez-Berrocal FJ, Rodríguez-Girondo M, et al. A new approach to epigenome-wide discovery of non-invasive methylation biomarkers for colorectal cancer screening in circulating cell-free DNA using pooled samples. Clinical epigenetics. 2018;10(1):53.

11. Salameh Y, Bejaoui Y, El Hajj N. DNA methylation biomarkers in aging and age-related diseases. Frontiers in Genetics. 2020;11:171.

12. Mendizabal I, Berto S, Usui N, Toriumi K, Chatterjee P, Douglas C, et al. Cell type-specific epigenetic links to schizophrenia risk in the brain. Genome biology. 2019;20(1):135.

13. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008;454(7205):766-70.

14. Doi A, Park I-H, Wen B, Murakami P, Aryee MJ, Irizarry R, et al. Differential methylation of tissue-and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nature genetics. 2009;41(12):1350.

15. Chen D-P, Lin Y-C, Fann CS. Methods for identifying differentially methylated regions for sequence-and array-based data. Briefings in functional genomics. 2016;15(6):485-90.

16. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC bioinformatics. 2012;13(1):86.

17. Lin X, Teh AL, Chen L, Lim IY, Tan PF, MacIsaac JL, et al. Choice of surrogate tissue influences neonatal EWAS findings. BMC medicine. 2017;15(1):211.

18. Miller S, Dykes D, Polesky H. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic acids research. 1988;16(3):1215.

19. Jaffe AE, Murakami P, Lee H, Leek JT, Fallin MD, Feinberg AP, et al. Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. International journal of epidemiology. 2012;41(1):200-9.

20. Shaffer JP. Multiple hypothesis testing. Annual review of psychology. 1995;46(1):561-84.

21. Butcher LM, Beck S. Probe Lasso: a novel method to rope in differentially methylated regions with 450K DNA methylation data. Methods. 2015;72:21-8.

22. Satterthwaite FE. An approximate distribution of estimates of variance components. Biometrics bulletin. 1946;2(6):110-4.

23. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, Lord RV, et al. De novo identification of differentially methylated regions in the human genome. Epigenetics & chromatin. 2015;8(1):6.

24. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC bioinformatics. 2011;12(1):35.

25. Wickham H. ggplot2: elegant graphics for data analysis: Springer; 2016.

26. Team RC. R: a language and environment for statistical computing, version 3.0. 2. Vienna, Austria: R Foundation for Statistical Computing; 2013. 2019.

27. Eckert MA, Coscia F, Chryplewicz A, Chang JW, Hernandez KM, Pan S, et al. Proteomics reveals NNMT as a master metabolic regulator of cancer-associated fibroblasts. Nature. 2019;569(7758):723-8.

28. Mallik S, Odom GJ, Gao Z, Gomez L, Chen X, Wang L. An evaluation of supervised methods for identifying differentially methylated regions in Illumina methylation arrays. Briefings in bioinformatics. 2019;20(6):2224-35.

29. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-9.

30. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome biology. 2016;17(1):208.

31. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. Nucleic acids research. 2015;44(3):e20-e.

32. Xu Z, Langie SA, De Boever P, Taylor JA, Niu L. RELIC: a novel dye-bias correction method for Illumina Methylation BeadChip. BMC genomics. 2017;18(1):4.

33. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2012;29(2):189-96.

34. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011;98(4):288-95.

35. Eckmann-Scholz C, Bens S, Kolarova J, Schneppenheim S, Caliebe A, Heidemann S, et al. DNA-methylation profiling of fetal tissues reveals marked epigenetic differences between chorionic and amniotic samples. PLoS One. 2012;7(6):e39014.

36. Poggi N. Microbenchmark. In: Sakr S, Zomaya AY, editors. Encyclopedia of Big Data Technologies. Cham: Springer International Publishing; 2019. p. 1143-52.

37. Sofer T, Schifano ED, Hoppin JA, Hou L, Baccarelli AA. A-clustering: a novel method for the detection of co-regulated methylation regions, and regions associated with exposure. Bioinformatics. 2013;29(22):2884-91.

38. Teschendorff AE, Relton CL. Statistical and integrative system-level analysis of DNA methylation data. Nature Reviews Genetics. 2018;19(3):129.

39. Ruiz-Arenas C, González JR. Redundancy analysis allows improved detection of methylation changes in large genomic regions. BMC bioinformatics. 2017;18(1):553.

Statement of Authorship

Title of Paper	Comparison of methods for region-based differential methylation analysis for Illumina EPIC arrays
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Qianhui Wan		
Contribution to the Paper	Performed the data analyses and wrote the manuscript.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	23/03/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Shalem Yiner-Lee Leemaqz		
Contribution to the Paper	Assisted with statistics and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Stephen Martin Pederson		
Contribution to the Paper	Assisted with statistics and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Tanja Jankovic-Karasoulos		
Contribution to the Paper	Laboratory work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Dale Christopher McAninch		
Contribution to the Paper	Laboratory work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Dylan McCullough		
Contribution to the Paper	Sample collection and laboratory work.		
Signature		Date	23/03/20

Name of Co-Author	Melanie Denise Smith		
Contribution to the Paper	Assisted in coding and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Tina Bianco-Miotto		
Contribution to the Paper	Supervised the work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Claire Trelford Roberts		
Contribution to the Paper	Supervised the work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	James Breen		
Contribution to the Paper	Conceived the study design, supervised the work, assisted in coding and edited the manuscript.		
Signature		Date	23/03/20

Please cut and paste additional co-author panels here as required.

3 Quality control measures for placental sample purity in DNA methylation array analyses

3.1 Manuscript

Placenta 88 (2019) 8–11

Contents lists available at ScienceDirect

Placenta

journal homepage: www.elsevier.com/locate/placenta

Technical note

Quality control measures for placental sample purity in DNA methylation array analyses

Qianhui Wan[a,b], Shalem Yiner-Lee Leemaqz[a,b], Stephen Martin Pederson[c], Dylan McCullough[a,b], Dale Christopher McAninch[a,b], Tanja Jankovic-Karasoulos[a,b], Melanie Denise Smith[a,b], Konstantinos Justinian Bogias[a,b], Ning Liu[b,c], James Breen[a,c,d], Claire Trelford Roberts[a,b], Tina Bianco-Miotto[a,e,*]

[a] Robinson Research Institute, University of Adelaide, Adelaide, SA, Australia
[b] Adelaide Medical School, University of Adelaide, Adelaide, SA, Australia
[c] Bioinformatics Hub, School of Biological Sciences, University of Adelaide, Adelaide, SA, Australia
[d] South Australian Health & Medical Research Institute (SAHMRI), Adelaide, SA, Australia
[e] School of Agriculture, Food and Wine, University of Adelaide, Adelaide, SA, Australia

ARTICLE INFO

Keywords:
Human placenta
DNA methylation
Methylation Bead Chip
Quality control

ABSTRACT

The purity of tissue samples can affect the accuracy and utility of DNA methylation array analyses. This is particularly important for the placenta which is globally hypomethylated compared to other tissues. Placental villous tissue from early pregnancy terminations can be difficult to separate from non-villous tissue, resulting in potentially inaccurate results. We used several methods to identify mixed placenta samples using DNA methylation array datasets from our laboratory and those contained in the NCBI GEO database, highlighting the importance of determining sample purity during quality control processes.

1. Introduction

The number of studies investigating genome-scale DNA methylation is growing rapidly. The most popular platforms are DNA methylation arrays [1]. The placenta is a transient organ that primarily acts to support fetal differentiation and growth and orchestrates maternal adaptations to pregnancy [2,3]. Placental DNA is uniquely and globally hypomethylated, reflecting this organ's early developmental origin and distinct functions [4,5]. Analysing the placental methylome could provide information relevant to understanding pregnancy health and disease [6]. Existing pipelines for processing DNA methylation array data do not include a quality control step for detecting tissue impurities [7,8]. We demonstrate that mixed placenta samples, that is placenta samples that contain a significant amount of surrounding contaminating tissue, can be identified from DNA methylation array data, allowing for sample removal or incorporation of other analytic changes as appropriate.

2. Methods

2.1. Array data

Thirteen human tissue datasets containing raw IDAT files from Illumina Infinium HumanMethylation450 (450 K) or HumanMethylationEPIC (EPIC) array platforms (GPL13534 or GPL21145) were acquired from NCBI GEO database and our laboratory. Ten datasets used the 450 K platform (GSE66210, GSE74738, GSE69502, GSE75196, GSE75248, GSE120250, GSE100197, GSE98224, GSE71678, GSE66459) [9–18] and 3 the EPIC platform (GSE115508, GSE113600 [19,20] plus 10 samples from our study, GSE131945). 410 samples (394 from 450 K platform and 16 from EPIC platform) were selected: 380 placentas [22 from first trimester, 16 from second trimester (pregnancy terminations) and 342 from term (uncomplicated pregnancies)], 12 maternal whole blood samples, 11 umbilical cord blood, 3 decidua, 2 amnion and 2 chorion (Supplementary Table 1). For the term samples, only samples from uncomplicated pregnancies were used in the downstream analyses.

* Corresponding author. Robinson Research Institute, University of Adelaide, Adelaide, SA, Australia.
E-mail address: tina.bianco@adelaide.edu.au (T. Bianco-Miotto).

https://doi.org/10.1016/j.placenta.2019.09.006
Received 12 June 2019; Received in revised form 3 September 2019; Accepted 15 September 2019

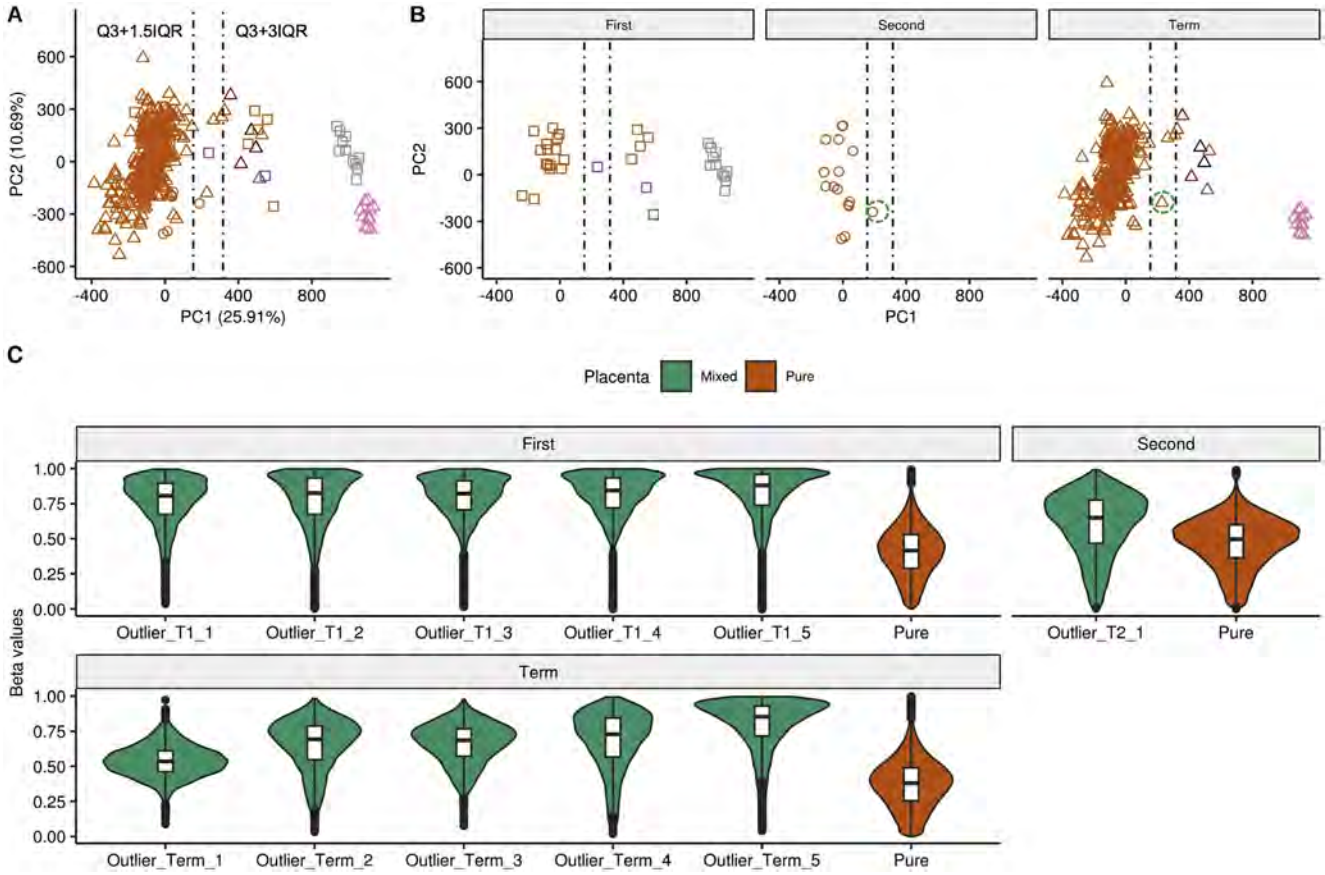

Fig. 1. Analysis of 408 samples from different tissue types. (A) PCA plot using M values for all 408 samples. Trimester: □first, ○second, △term. Tissue types: placenta (orange), amnion (black), chorion (red), maternal whole blood (grey), decidua (purple), umbilical cord blood (pink). Dashed vertical line marked Q3 + 1.5IQR (153.88) and Q3 + 3IQR (316.01) of PC1. Outliers are at the right side of dashed vertical line (Q3 + 1.5IQR). (B) PCA plot of samples from first trimester, second trimester and term were shown respectively. Mixed placenta samples tend to be similar to non-placenta tissues. Samples in dashed green ellipses were Outlier_T2_1 and Outlier_Term_1. (C) DNA methylation difference between outlier and pure placenta from first trimester, second trimester and term. T1: first trimester; T2: second trimester.

2.2. Array quality control

Two samples were removed due to low array intensities defined as < 10.5, one first trimester placenta sample (GSM1617002) and one maternal whole blood sample (GSM1616993). Probes filtered out included: failed probes with detection P > 0.01, probes with fewer than 3 beads in > 5% of all samples [8]; cross-reactive probes [21]; probes on sex chromosomes and probes with biased DNA methylation signal due to SNPs at CpG sites, single base extension (SBE) sites and probe body [22,23]. In all, 408 samples and 301,496 probes common to both 450 K and EPIC arrays were used for downstream analyses. See Supplementary Methods for additional information.

2.3. Principal component analysis (PCA) and sample clustering

The dye bias correction for filtered data was performed using the method regression on logarithm of internal control probes (RELIC) from *ENmix* package and background subtraction according to tissue types were performed with the method from *ENmix* package which models methylation signal intensities with a flexible exponential-normal mixture distribution, and use out-of-band Infinium I intensities ("OOB") to estimate normal distribution parameters to model background noise. Datasets were normalised using the beta-mixture quantile normalization method [24] to correct for type I and type II probe bias. Control probes were used to check batch effects between studies. All 408 samples were used to generate the PCA plot which gave an overview of similarities and differences between samples from different tissue types. Interquartile range (IQR) of values for samples at PC1 was calculated to estimate outliers. Using PC1, placenta samples with less than Q3 + 1.5IQR were considered pure placenta samples and those samples with greater than Q3 + 3IQR were mixed placenta samples. The area in between Q3 + 1.5IQR and Q3 + 3IQR was designated a 'grey' zone. Multivariate unsupervised clustering with mixtures of Gaussian distributions [25] was performed to estimate the probability of each sample being a pure placenta sample. The top 2% of probes that were different between pure and mixed samples (PC1) were selected to investigate DNA methylation differences between them. M values were used for the statistical tests based on linear models.

2.4. DNA methylation analyses at partially methylated domains (PMDs) and placenta-specific imprinting control regions (ICRs)

The genome was tailed by 10 kb non-overlapping bins and mean Beta values (methylation percentage) of each bin were calculated [5]. Mean Beta values of bins that overlapped with PMDs for pure and mixed placenta samples were plotted to show the difference between these samples. Beta values at ICRs [26] for pure and mixed samples were also plotted. Mann–Whitney *U* test was used to test the significance of change of DNA methylation at PMDs and ICRs between potentially mixed and pure placenta sample groups. The codes for all analyses can be found here: https://github.com/QianhuiWan/MethylationArray_Placenta.

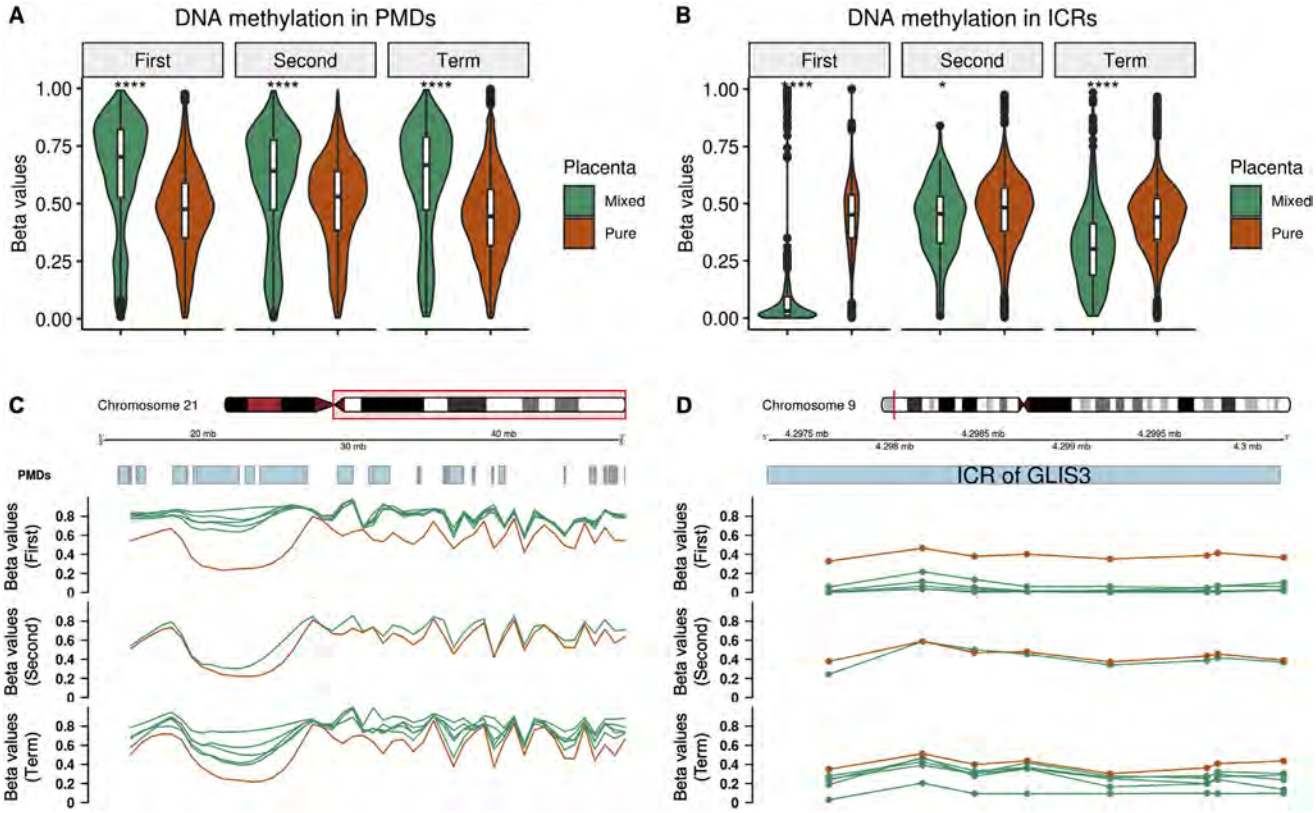

Fig. 2. DNA methylation in partially methylated domains (PMDs) (A and C) and imprinting control regions (ICRs) (B and D). (A) Average DNA methylation in PMDs was higher in mixed (green) than pure placenta samples (orange) ($p < 2.2e\text{-}16$), Chr21 was used as an example as shown in (C). (B) DNA methylation at placenta-specific ICRs was overall lower in mixed (green) than in pure placenta samples (orange) ($p < 0.05$), ICR on Chr9 is an example shown in (D). Green lines in (C) and (D) represent individual mixed samples from different trimesters, while orange lines in (C) and (D) represent mean beta values of pure samples from different trimesters. Mann–Whitney *U* test was used to test the significance between groups in plot (A) and (B). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 2.2e\text{-}16$.

3. Results

The PCA plot of 408 samples identified 11 placenta samples as outliers (Fig. 1A). Clustering of a subset of placental samples to non-placental samples suggests that these samples may be mixed with DNA from non-placenta tissue. These samples will be called “mixed placenta samples” in this report (Supplementary Fig. 1). Five were from first trimester, one from second trimester and the remaining five were from term samples (Fig. 1B). Four of the term outliers were previously identified [14]. Nine placenta sample outliers clustered with decidua, amnion and chorion and are likely to contain more non-placenta tissue than the other two samples (Outlier_T2_1 and Outlier_Term_1, dashed green circle; Fig. 1B). The estimated probability of each placenta sample being a pure sample is listed in Supplementary Table 2. The top 2% of most variable probes according to PC1, showed that all had higher DNA methylation in the putative mixed placenta samples compared to the remaining placenta samples (Fig. 1C).

Mixed placenta samples from first trimester, second trimester and term, showed greater DNA methylation than pure placenta samples at PMDs ($P < 2.2e\text{-}16$) (Fig. 2A and C). Mixed placenta samples showed altered DNA methylation compared to pure placenta samples at placenta-specific ICRs ($P < 0.05$) (Fig. 2B and D and Supplementary Fig. 2).

4. Discussion

Using publicly available DNA methylation array datasets, as well as 10 samples of our own, we identified outlier placenta villous tissue samples which clustered with other tissue types and were likely to contain other tissues, most likely decidua. These outliers had different DNA methylation profiles at both PMDs and ICRs, indicating they may contain non-villous tissue, such as hypermethylated maternal decidua [27]. Placenta samples from early gestation are at particular risk of being mixed with other tissue types, possibly because the termination procedures inevitably macerate the tissue. Since these outliers had quite distinct methylation profiles compared to other placenta samples, they could potentially influence results, necessitating their removal or down weighting before further analysis.

Before processing high dimension data from placenta villous tissues, we recommend checking sample purity in three steps. First, download the DNA methylation array data listed in Supplementary Table 1 from NCBI GEO database. Second, apply PCA and unsupervised clustering using pure placenta samples from GEO database as positive controls and samples from other tissue types from GEO database as negative controls. Third, if any samples cluster with other tissue types, and the estimated probability of the sample being pure placenta is low, the DNA methylation of placenta PMDs and ICRs needs to be checked to verify whether these samples should be removed or down weighted.

Conflicts of interest

The authors confirm that there are no conflicts of interest.

Acknowledgments

We would like to thank all the women who have kindly donated their placentas for our research and the clinical staff who collected them.

CTR was supported by a Lloyd Cox Professorial Research Fellowship from the University of Adelaide. The research was funded by a NICHD

NIH Project Grant (NIH oppRFA-HD-16-036 1 R01 HD089685-01): Maternal molecular profiles reflect placental function and development across gestation awarded to CTR, TBM and JB.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.placenta.2019.09.006.

References

[1] A.E. Teschendorff, C.L. Relton, Statistical and integrative system-level analysis of DNA methylation data, Nat. Rev. Genet. 19 (2017) 129–147, https://doi.org/10.1038/nrg.2017.86.

[2] N.M. Gude, C.T. Roberts, B. Kalionis, R.G. King, Growth and function of the normal human placenta, Thromb. Res. 114 (2004) 397–407, https://doi.org/10.1016/j.thromres.2004.06.038.

[3] S. Zhang, T. Regnault, P. Barker, K. Botting, I. McMillen, C. McMillan, C. Roberts, J. Morrison, S. Zhang, T.R.H. Regnault, P.L. Barker, K.J. Botting, I.C. McMillen, C.M. McMillan, C.T. Roberts, J.L. Morrison, Placental adaptations in growth restriction, Nutrients 7 (2015) 360–389, https://doi.org/10.3390/nu7010360.

[4] T. Bianco-Miotto, B.T. Mayne, S. Buckberry, J. Breen, C.M. Rodriguez Lopez, C.T. Roberts, Recent progress towards understanding the role of DNA methylation in human placental development, Reproduction 152 (2016) R23–R30, https://doi.org/10.1530/REP-16-0014.

[5] D.I. Schroeder, J.D. Blair, P. Lott, H.O.K. Yu, D. Hong, F. Crary, P. Ashwood, C. Walker, I. Korf, W.P. Robinson, J.M. LaSalle, The human placenta methylome, Proc. Natl. Acad. Sci. 110 (2013) 6037–6042, https://doi.org/10.1073/PNAS.1215145110.

[6] W.P. Robinson, E.M. Price, The human placental methylome, Cold Spring Harb. Perspect. Med. 5 (2015) a023044, https://doi.org/10.1101/cshperspect.a023044.

[7] J.-P. Fortin, T.J. Triche, K.D. Hansen, Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi, Bioinformatics 33 (2016) btw691, https://doi.org/10.1093/bioinformatics/btw691.

[8] Y. Tian, T.J. Morris, A.P. Webster, Z. Yang, S. Beck, A. Feber, A.E. Teschendorff, ChAMP: updated methylation analysis pipeline for Illumina Bead Chips, Bioinformatics 33 (2017) 3982–3984, https://doi.org/10.1093/bioinformatics/btx513.

[9] L. Hatt, M.M. Aagaard, J. Graakjaer, C. Bach, S. Sommer, I.E. Agerholm, S. Kølvraa, A. Bojesen, Microarray-based analysis of methylation status of CpGs in placental DNA and maternal blood DNA – potential new epigenetic biomarkers for cell free fetal DNA-based diagnosis, PLoS One 10 (2015) e0128918, , https://doi.org/10.1371/journal.pone.0128918.

[10] C.W. Hanna, M.S. Peñaherrera, H. Saadeh, S. Andrews, D.E. McFadden, G. Kelsey, W.P. Robinson, Pervasive polymorphic imprinted methylation in the human placenta, Genome Res. 26 (2016) 756–767, https://doi.org/10.1101/gr.196139.115.

[11] E.M. Price, M.S. Peñaherrera, E. Portales-Casamar, P. Pavlidis, M.I. Van Allen, D.E. McFadden, W.P. Robinson, Profiling placental and fetal DNA methylation in human neural tube defects, Epigenet. Chromatin 9 (2016) 6, https://doi.org/10.1186/s13072-016-0054-8.

[12] K.R. Yeung, C.L. Chiu, R. Pidsley, A. Makris, A. Hennessy, J.M. Lind, DNA methylation profiles in preeclampsia and healthy control placentas, Am. J. Physiol. Cell Physiol. 310 (2016) H1295–H1303, https://doi.org/10.1152/ajpheart.00958.2015.

[13] A.G. Paquette, E.A. Houseman, B.B. Green, C. Lesseur, D.A. Armstrong, B. Lester, C.J. Marsit, Regions of variable DNA methylation in human placenta associated with newborn neurobehavior, Epigenetics 11 (2016) 603–613, https://doi.org/10.1080/15592294.2016.1195534.

[14] S. Choufani, A.L. Turinsky, N. Melamed, E. Greenblatt, M. Brudno, A. Bérard, W.D. Fraser, R. Weksberg, J. Trasler, P. Monnier, For the 3D cohort study group, impact of assisted reproduction, infertility, sex, and paternal factors on the placental DNA methylome, Hum. Mol. Genet. (2018), https://doi.org/10.1093/hmg/ddy321.

[15] S.L. Wilson, K. Leavey, B.J. Cox, W.P. Robinson, Mining DNA methylation alterations towards a classification of placental pathologies, Hum. Mol. Genet. 27 (2018) 135–146, https://doi.org/10.1093/hmg/ddx391.

[16] K. Leavey, S.L. Wilson, S.A. Bainbridge, W.P. Robinson, B.J. Cox, Epigenetic regulation of placental gene expression in transcriptional subtypes of preeclampsia, Clin. Epigenet. 10 (2018) 28, https://doi.org/10.1186/s13148-018-0463-6.

[17] B.B. Green, M.R. Karagas, T. Punshon, B.P. Jackson, D.J. Robbins, E.A. Houseman, C.J. Marsit, Epigenome-Wide assessment of DNA methylation in the placenta and arsenic exposure in the New Hampshire birth cohort study (USA), Environ. Health Perspect. 124 (2016) 1253–1260, https://doi.org/10.1289/ehp.1510437.

[18] F. Fernando, R. Keijser, P. Henneman, A.-M.F. van der Kevie-Kersemaekers, M.M. Mannens, J.A. van der Post, G.B. Afink, C. Ris-Stalpers, The idiopathic preterm delivery methylation profile in umbilical cord blood DNA, BMC Genomics 16 (2015) 736, https://doi.org/10.1186/s12864-015-1915-4.

[19] C. Konwar, E.M. Price, L.Q. Wang, S.L. Wilson, J. Terry, W.P. Robinson, DNA methylation profiling of acute chorioamnionitis-associated placentas and fetal membranes: insights into epigenetic variation in spontaneous preterm births, Epigenet. Chromatin 11 (2018) 63, https://doi.org/10.1186/s13072-018-0234-9.

[20] M. Yu, G. Du, Q. Xu, Z. Huang, X. Huang, Y. Qin, L. Han, Y. Fan, Y. Zhang, X. Han, Z. Jiang, Y. Xia, X. Wang, C. Lu, Integrated analysis of DNA methylome and transcriptome identified CREB5 as a novel risk gene contributing to recurrent pregnancy loss, EBioMedicine 35 (2018) 334–344, https://doi.org/10.1016/j.ebiom.2018.07.042.

[21] R. Pidsley, E. Zotenko, T.J. Peters, M.G. Lawrence, G.P. Risbridger, P. Molloy, S. Van Djik, B. Muhlhausler, C. Stirzaker, S.J. Clark, Critical evaluation of the Illumina Methylation EPIC BeadChip microarray for whole-genome DNA methylation profiling, Genome Biol. (2016), https://doi.org/10.1186/s13059-016-1066-1.

[22] D. Zhi, S. Aslibekyan, M.R. Irvin, S.A. Claas, I.B. Borecki, J.M. Ordovas, D.M. Absher, D.K. Arnett, SNPs located at CpG sites modulate genome-epigenome interaction, Epigenetics 8 (2013) 802–806, https://doi.org/10.4161/epi.25501.

[23] W. Zhou, P.W. Laird, H. Shen, Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes, Nucleic Acids Res. (2017), https://doi.org/10.1093/nar/gkw967.

[24] S. Horvath, DNA methylation age of human tissues and cell types, Genome Biol. 14 (2013) R115, https://doi.org/10.1186/gb-2013-14-10-r115.

[25] L. Scrucca, M. Fop, T.B. Murphy, A.E. Raftery, Mclust 5: clustering, classification and density estimation using Gaussian finite mixture models, R J. 8 (2016) 289–317 http://www.ncbi.nlm.nih.gov/pubmed/27818791 , Accessed date: 23 May 2019.

[26] F. Court, C. Tayama, V. Romanelli, A. Martin-Trujillo, I. Iglesias-Platas, K. Okamura, N. Sugahara, C. Simón, H. Moore, J.V. Harness, H. Keirstead, J.V. Sanchez-Mut, E. Kaneki, P. Lapunzina, H. Soejima, N. Wake, M. Esteller, T. Ogata, K. Hata, K. Nakabayashi, D. Monk, Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment, Genome Res. 24 (2014) 554–569, https://doi.org/10.1101/gr.164913.113.

[27] T.J. Jensen, S.K. Kim, Z. Zhu, C. Chin, C. Gebhard, T. Lu, C. Deciu, D. van den Boom, M. Ehrich, Whole genome bisulfite sequencing of cell-free DNA and its cellular contributors uncovers placenta hypomethylated domains, Genome Biol. 16 (2015) 78, https://doi.org/10.1186/s13059-015-0645-x.

3.2 Supplementary information

3.2.1 Supplementary methods

3.2.1.1 Sample and probe filtering

Unwanted samples and probes were first filtered out from datasets including samples from complicated pregnancies, samples with low quality and unwanted probes. Samples with median DNA methylation or unmethylation intensities <10.5 were excluded [1]. We also filtered failed probes with detection P>0.01, probes with fewer than 3 beads in >5% of all samples [2]; cross-reactive probes [3]; probes on sex chromosomes and probes with biased DNA methylation signal due to SNPs at CpG sites and single base extension (SBE) sites [4]. In favour of downstream analyses, the probes on X/Y chromosomes were removed to reduce sex bias and probes with SNPs within probe body (not located at CpG/SBE sites) were also removed [5]. In all, 408 samples and 310,882 probes common across all samples from both 450K and EPIC arrays were used for downstream analyses.

3.2.1.2 Principal component analysis

Since detection sensitivity is delta beta greater than 0.2 (95% confidence level) across greater than 90% of the loci for any given pair of samples [6], the top probes which were significantly correlated with PC1 and have absolute beta value differences > 0.2 [6, 7] between pure and mixed placenta samples were selected to investigate whether DNA methylation is increased or decreased in mixed samples. There were 6217 probes in the top 2% probes correlated with PC1, 6071 of these probes have absolute

beta value differences > 0.2. In these 6217 probes, 1971 probes were located in PMDs (in total 45311 probes in PMDs) and 3 probes located in placenta specific ICRs (in total 630 probes in ICRs).

3.2.1.3 Sample quality control with ewastools R package

Methods from *ewastools* package were used to check the homogeneity of genotypes (using SNP probes) for each sample [8]. The theory of genotype checking is one sample (or one person/individual) could only have 3 status of DNA methylation (MM or MU or UU) for a particular SNP, so the density plot of SNP methylation should be a plot with 3 peaks. If there are many SNP probes with methylation between 0%-50% or 50%-100% for one sample, this sample is likely to have more technical variance (or potential contamination from other individuals or other tissue types) than other samples that have 3 peaks as expected. (MM: both alleles were methylated (peak at 100% methylation), often referred to as AA or BB genotype; MU: one allele methylated, one allele unmethylated (peak at 50% methylation), often referred to as A/B genotype; UU: both alleles were unmethylated (peak at 0% methylation), often referred to as AA or BB genotype.)

3.2.1.4 Other statistical analyses

ANOVA *F*-statistic was used to test the contribution of tissue type, trimester, sex and batch groups to PC1/PC2 [9]. Mann–Whitney U test was used to test the significance of DNA methylation changes at PMDs and ICRs between potentially mixed and pure placenta sample groups [10].

3.2.2 Supplementary figures and tables

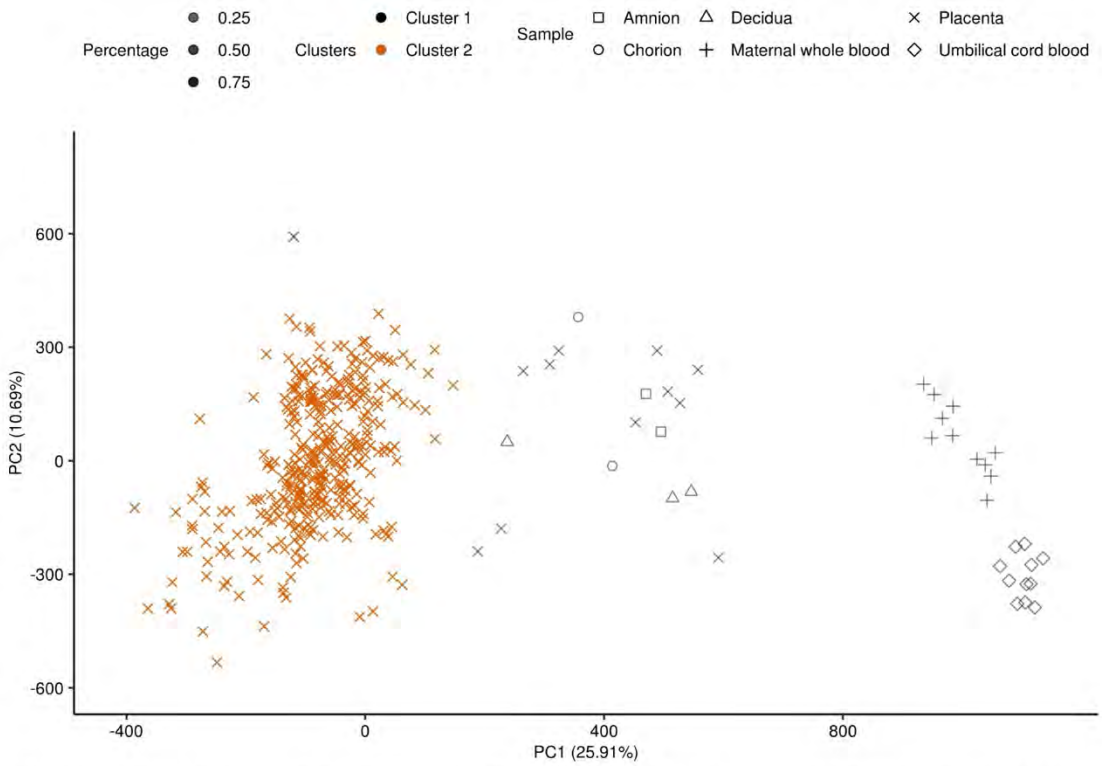

Figure S3.1 Identification of mixed placenta samples. Clusters identified by multivariate unsupervised clustering showed 11 mixed placenta samples were clustered with other tissue types instead of placenta. The probability that a sample in the data belongs to Cluster 2 was used to estimate the probability of each sample being a pure placenta sample. Percentage of placenta for each sample were labelled by transparency of the colour.

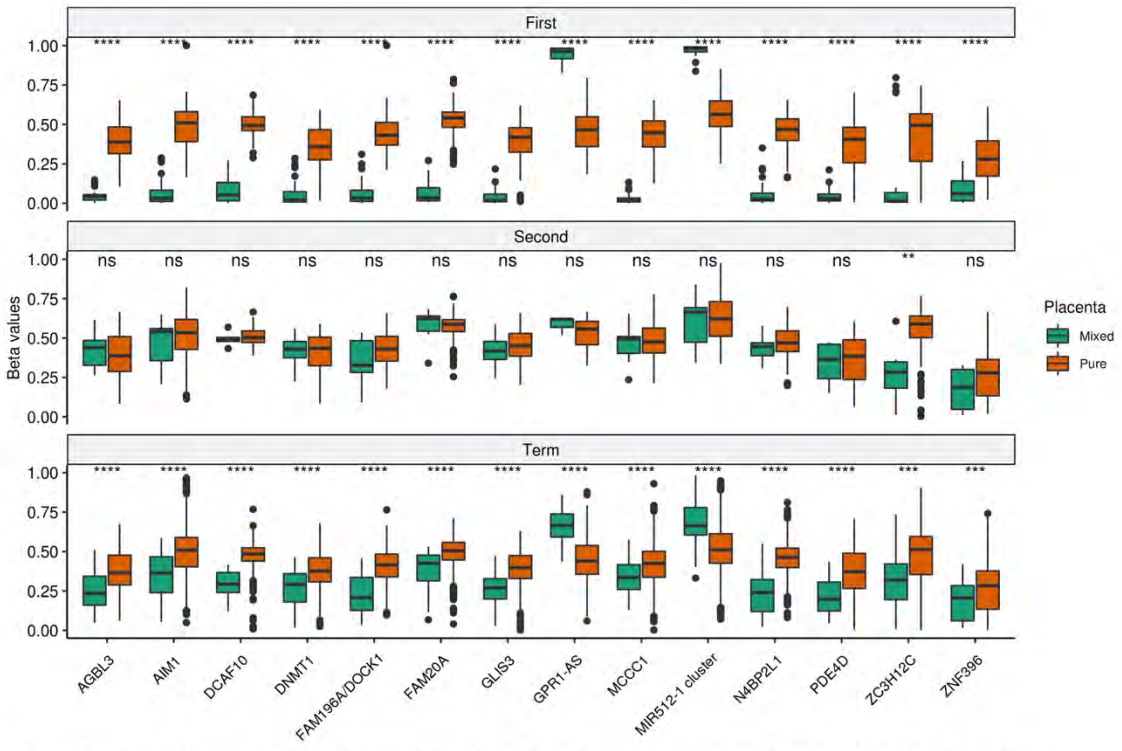

Figure S3.2 DNA methylation of all the placenta specific imprinting control regions (ICRs) for placenta samples. Genes that were controlled by corresponding ICRs were labelled on x axis. Beta values represent the average percentage of DNA methylation for ICRs. Mixed placenta samples were not imprinted like pure placenta samples. Mann-Whitney *U* test were used to test the significance between groups. **P*<0.05, ***P*<0.01, ****P*<0.001, *****P*<2.2e-16.

Table S3.1 DNA methylation datasets containing placental tissue samples (n=410) used in this study.

NO.	GEO accession	Platform	Number of samples	Tissue types	Gestational age range (weeks/trimester)	Ref. in main text
1	GSE66210	450K	24	Maternal blood cell samples from uncomplicated singleton pregnancies (n=12); CVS samples from uncomplicated pregnancies (n=12).	First trimester	Lotte et al. [9]
2	GSE74738	450K	27	Term placental chorionic villous samples from uncomplicated pregnancies (n=22); Extra-embryonic cell types from uncomplicated pregnancies, including amnion (n=2), chorion (n=2), and maternal decidua (n=1).	37.1-41.7	Courtney et al. [10]
3	GSE69502	450K	16	Placental chorionic villous samples from terminated pregnancies. Exclusion criteria for control cases included chromosomal abnormality, congenital or brain abnormality, or grossly abnormal placenta.	14.5-23.9	Price et al. [11]
4	GSE75196	450K	16	Term placenta samples from uncomplicated pregnancies. Placental biopsies were collected from five sites from the fetal side of the placenta.	38-40	Yeung et al. [12]
5	GSE75248	450K	174	Placenta samples from uncomplicated pregnancies. All samples were taken from the maternal side of the placenta, 2 cm from the umbilical cord insertion site, free of maternal decidua (Rhode Island Child Health Study (RICHS)).	Term	Paquette et al. [13]
6	GSE120250	450K	44	Term placenta samples from uncomplicated singleton pregnancies. Biopsies taken 1 cm below the chorionic plate were used for this study.	38.7±1.5	Choufani et al. [14]
7	GSE100197	450K	19	Term placental chorionic villous samples from uncomplicated pregnancies.	37-40	Wilson et al. [15]
8	GSE98224	450K	9	Term placenta samples from uncomplicated pregnancies (n=9). Placental tissue biopsies (1.5 x 1.5 cm cores through the full thickness of the placenta, excluding chorionic plate) are collected from a site within each quadrant, avoiding areas with obvious evidence of thrombosis or other abnormalities when possible (RCWIH BioBank).	38-40	Leavey et al. [16]; Wilson et al. [15]
9	GSE71678	450K	54	Term placenta samples from uncomplicated pregnancies. Placenta was biopsied adjacent to the cord insertion, removing any maternal decidua.	37.3-42.1	Green et al. [17]

10	GSE66459	450K	11	Term umbilical cord blood samples from uncomplicated pregnancies.	38-41.6	Fernando et al. [18]
11	GSE115508	EPIC	4	Term placental chorionic villous samples from uncomplicated pregnancies.	37	Konwar et al. [19]
12	GSE113600	EPIC	2	Maternal decidual samples from terminated pregnancies.	Mean gestational age = 7.083	Yu et al. [20]
13	GSE131945	EPIC	10	Placental chorionic villous samples from terminated pregnancies.	6-9	Our study

† CVS: chorionic villus sampling; 450K: Illumina Infinium HumanMethylation450 BeadChip; EPIC: Illumina Infinium HumanMethylationEPIC BeadChip.

Table S3.2 DNA methylation datasets containing placental tissue samples (n=410) used in this study.

Sample name	Tissue types	PC1	PC2	Trimester	Fetal sex	Study	Outlier	Pure placenta probability (Mclust)
GSM1622785_Term_8	Umbilical cord blood	1099.64256	-380.41329	Term	Female	GSE66459	NotPlacenta	3.74E-60
GSM1622782_Term_5	Umbilical cord blood	1111.07156	-283.16744	Term	Male	GSE66459	NotPlacenta	1.27E-58
GSM1622786_Term_9	Umbilical cord blood	1086.55004	-335.16342	Term	Female	GSE66459	NotPlacenta	1.01E-57
GSM1622779_Term_1	Umbilical cord blood	1068.90426	-370.14242	Term	Male	GSE66459	NotPlacenta	5.21E-57
GSM1622788_Term_11	Umbilical cord blood	1074.80093	-328.82147	Term	Female	GSE66459	NotPlacenta	1.79E-56
GSM1622787_Term_10	Umbilical cord blood	1074.05606	-287.81829	Term	Female	GSE66459	NotPlacenta	2.40E-55
GSM1622783_Term_6	Umbilical cord blood	1051.31277	-348.53342	Term	Male	GSE66459	NotPlacenta	7.65E-55
GSM1622784_Term_7	Umbilical cord blood	1078.32086	-236.66873	Term	Male	GSE66459	NotPlacenta	2.00E-54
GSM1622789_Term_12	Umbilical cord blood	1052.78322	-312.43556	Term	Male	GSE66459	NotPlacenta	4.69E-54
GSM1622780_Term_2	Umbilical cord blood	1048.27929	-257.20596	Term	Female	GSE66459	NotPlacenta	2.87E-52
GSM1622781_Term_3	Umbilical cord blood	1038.89688	-270.49842	Term	Male	GSE66459	NotPlacenta	8.99E-52
GSM1616987_8795207029_R06C01	Maternal whole blood	981.397796	-115.60884	First	Female	GSE66210	NotPlacenta	2.58E-43
GSM1616986_8795207029_R04C01	Maternal whole blood	1004.19263	-25.204566	First	Female	GSE66210	NotPlacenta	3.83E-43
GSM1616992_7668610068_R05C02	Maternal whole blood	981.534471	11.0970071	First	Female	GSE66210	NotPlacenta	1.39E-40
GSM1616989_8795207029_R04C02	Maternal whole blood	976.880022	1.47491118	First	Female	GSE66210	NotPlacenta	1.99E-40
GSM1616991_7668610068_R04C02	Maternal whole blood	957.884367	11.2502587	First	Female	GSE66210	NotPlacenta	8.97E-39
GSM1616997_8795194156_R04C02	Maternal whole blood	963.088232	90.070672	First	Female	GSE66210	NotPlacenta	1.47E-37

GSM1616995_8795194156_R01C02	Maternal whole blood	940.088401	87.4688221	First	Female	GSE66210	NotPlacenta	6.34E-36
GSM1616996_8795194156_R04C01	Maternal whole blood	924.125899	60.6413811	First	Female	GSE66210	NotPlacenta	2.71E-35
GSM1616990_7668610068_R03C02	Maternal whole blood	937.818964	149.643496	First	Female	GSE66210	NotPlacenta	1.46E-34
GSM1616988_8795207029_R01C02	Maternal whole blood	924.003717	172.396044	First	Female	GSE66210	NotPlacenta	3.65E-33
GSM1616994_8795194156_R01C01	Maternal whole blood	913.902087	197.882087	First	Female	GSE66210	NotPlacenta	5.29E-32
GSM1931537_6042308147_R06C01	Decidua	586.283027	-177.42053	Term	Female	GSE74738	NotPlacenta	7.54E-18
201414140060_R01C01	Placenta	560.356836	-241.65999	First	Female	GSE131945	Outlier_T1_5	2.66E-17
GSM3109408_decidua1	Decidua	511.952529	-128.19929	First	Female	GSE113600	NotPlacenta	1.92E-13
201414140063_R06C01	Placenta	554.915234	269.167585	First	Female	GSE131945	Outlier_T1_4	2.56E-11
GSM1931538_6042324125_R05C02	Amnion	500.091144	88.8520621	Term	Male	GSE74738	NotPlacenta	1.11E-10
201332340153_R03C01	Placenta	516.996928	199.630123	First	Female	GSE131945	Outlier_T1_3	2.22E-10
GSM3396872_9992576207_R02C02	Placenta	502.789755	138.806773	Term	Male	GSE120250	Outlier_Term_5	2.44E-10
GSM1931540_6042324158_R06C02	Amnion	476.224014	167.660338	Term	Male	GSE74738	NotPlacenta	4.92E-09
201414140063_R05C01	Placenta	485.161443	308.41095	First	Female	GSE131945	Outlier_T1_2	2.71E-08
201332340153_R01C01	Placenta	430.385117	145.203014	First	Female	GSE131945	Outlier_T1_1	1.69E-07
GSM1931542_6042324125_R06C02	Chorion	378.916294	-4.1580731	Term	Male	GSE74738	NotPlacenta	1.01E-06
GSM1931544_6042324158_R05C02	Chorion	398.206016	336.481448	Term	Male	GSE74738	NotPlacenta	2.73E-05
GSM3396859_9992571127_R03C02	Placenta	317.183037	297.000661	Term	Male	GSE120250	Outlier_Term_4	0.00281721
GSM3396835_9992571127_R06C02	Placenta	304.56366	268.62116	Term	Female	GSE120250	Outlier_Term_3	0.00460935
GSM1947213_6008581028_R01C01	Placenta	225.950963	-189.88882	Term	Female	GSE75248	Outlier_Term_1	0.00750981
GSM3396829_9992571130_R06C01	Placenta	255.362643	248.421983	Term	Female	GSE120250	Outlier_Term_2	0.05151119
GSM3109409_decidua2	Decidua	216.547463	-10.72306	First	Female	GSE113600	NotPlacenta	0.0725266
GSM1702177_7970368142_R05C02	Placenta	180.219273	-230.88243	Second	Female	GSE69502	Outlier_T2_1	0.08930114
GSM3179720_200889820023_R01C01	Placenta	-131.39561	564.481443	Term	Male	GSE115508	Pure	0.16679157
GSM3396869_9992576160_R03C02	Placenta	122.677331	287.542376	Term	Male	GSE120250	Pure	0.8942521
GSM3396839_9992576160_R05C01	Placenta	123.13862	219.282509	Term	Male	GSE120250	Pure	0.90526337
GSM3179755_200925700033_R04C01	Placenta	-146.56423	342.618981	Term	Male	GSE115508	Pure	0.91639743

GSM3396831_9992576159_R02C01	Placenta	106.031102	238.321292	Term	Male	GSE120250	Pure	0.93683523
GSM3396862_9992576022_R03C01	Placenta	111.404075	100.0095	Term	Female	GSE120250	Pure	0.93716079
GSM1617003_7668610068_R03C01	Placenta	-177.23783	290.267293	First	Male	GSE66210	Pure	0.93821973
GSM3179728_200889820024_R01C01	Placenta	-153.47057	313.206649	Term	Female	GSE115508	Pure	0.94210137
GSM1843078_3999984101_R01C02	Placenta	-280.44125	150.553103	Term	Male	GSE71678	Pure	0.94466312
GSM1702174_7970368142_R04C01	Placenta	98.9992954	174.870232	Second	Male	GSE69502	Pure	0.95387285
GSM3396864_9992576119_R04C01	Placenta	70.7398204	280.340661	Term	Male	GSE120250	Pure	0.96477681
GSM2589579_10005833024_R03C02	Placenta	-86.127585	339.414493	Term	Male	GSE98224	Pure	0.96613769
GSM3396854_9992576159_R04C02	Placenta	74.9294382	239.695504	Term	Male	GSE120250	Pure	0.9684213
GSM1947201_5806636027_R03C01	Placenta	20.149125	338.674255	Term	Female	GSE75248	Pure	0.97141428
GSM1947208_5806636027_R02C01	Placenta	46.093778	305.724964	Term	Male	GSE75248	Pure	0.97167267
GSM2589569_9977525015_R01C02	Placenta	-76.381831	332.484067	Term	Male	GSE98224	Pure	0.97180272
GSM3396843_9992571114_R01C01	Placenta	81.3121576	154.961098	Term	Female	GSE120250	Pure	0.97227072
GSM3396865_9992576022_R01C02	Placenta	74.0747637	138.549627	Term	Female	GSE120250	Pure	0.97788548
GSM3396871_9992576119_R05C01	Placenta	46.107981	267.489451	Term	Male	GSE120250	Pure	0.97794571
GSM1617005_7668610068_R01C02	Placenta	-37.55302	327.482794	First	Female	GSE66210	Pure	0.97801924
GSM1702240_9406922026_R04C01	Placenta	21.2287205	298.301688	Second	Male	GSE69502	Pure	0.97950928
GSM1702246_9406922117_R04C01	Placenta	12.5871547	306.356693	Second	Female	GSE69502	Pure	0.97957396
GSM3396837_9992576119_R01C02	Placenta	26.5543427	289.860139	Term	Female	GSE120250	Pure	0.97981029
GSM2589550_10005833037_R05C01	Placenta	35.7860513	271.280439	Term	Male	GSE98224	Pure	0.98037005
GSM3396846_9992576033_R02C01	Placenta	22.5704909	287.682639	Term	Female	GSE120250	Pure	0.98093305
GSM3396838_9992576119_R03C01	Placenta	30.7843734	262.24553	Term	Male	GSE120250	Pure	0.98269115
GSM1947101_6008581008_R06C01	Placenta	47.8952508	-372.99103	Term	Female	GSE75248	Pure	0.983
GSM3396836_9992576119_R05C02	Placenta	7.54822156	288.835395	Term	Male	GSE120250	Pure	0.98310232
GSM1947112_5806636034_R06C01	Placenta	48.4999943	184.921506	Term	Female	GSE75248	Pure	0.98536241
GSM3396860_9992576119_R04C02	Placenta	-23.74892	288.705687	Term	Female	GSE120250	Pure	0.98558333
GSM1931583_9296930103_R03C01	Placenta	-81.505228	275.211623	Term	Male	GSE74738	Pure	0.98651763

GSM2589571_10005833037_R03C02	Placenta	-17.903047	279.720663	Term	Female	GSE98224	Pure	0.98658354
GSM1617007_8795194156_R02C02	Placenta	6.97758692	257.512646	First	Female	GSE66210	Pure	0.98708958
GSM1842815_3999984081_R01C01	Placenta	61.0402927	19.1976045	Term	Male	GSE71678	Pure	0.98755547
GSM1947237_5859720011_R04C02	Placenta	60.2417358	16.8911165	Term	Male	GSE75248	Pure	0.98788198
GSM3396858_9992571114_R06C01	Placenta	43.2856375	160.646687	Term	Male	GSE120250	Pure	0.98818781
GSM1702222_9296930154_R03C02	Placenta	-4.7591189	257.368205	Second	Female	GSE69502	Pure	0.98828424
GSM1947152_5806636035_R06C01	Placenta	48.6754557	122.917215	Term	Male	GSE75248	Pure	0.98862079
GSM3396852_9992576159_R05C01	Placenta	8.65387748	239.587924	Term	Female	GSE120250	Pure	0.98862958
GSM2674510_9977525015_R03C02	Placenta	-185.80865	184.163163	Term	Female	GSE100197	Pure	0.98868435
GSM2589544_10005833110_R04C01	Placenta	-67.632767	265.587167	Term	Female	GSE98224	Pure	0.98878833
GSM2674511_10005833024_R02C01	Placenta	-53.472494	266.479923	Term	Male	GSE100197	Pure	0.98898435
GSM1843011_3999984005_R05C02	Placenta	-409.89184	-114.29458	Term	Male	GSE71678	Pure	0.98904094
GSM3396850_9992571127_R05C01	Placenta	22.9096248	207.475153	Term	Female	GSE120250	Pure	0.98922266
GSM3396855_9992576119_R03C02	Placenta	-21.854436	258.554475	Term	Female	GSE120250	Pure	0.98928672
GSM1931587_9285451020_R03C02	Placenta	-80.642264	251.303053	Term	Female	GSE74738	Pure	0.9901617
GSM1931590_9266441156_R03C01	Placenta	-92.171679	245.394109	Term	Female	GSE74738	Pure	0.99033498
GSM1947081_5806636079_R03C02	Placenta	-69.964035	250.88314	Term	Female	GSE75248	Pure	0.99057563
GSM1931588_9296930123_R04C02	Placenta	-72.108732	248.966376	Term	Female	GSE74738	Pure	0.9907389
201414140063_R08C01	Placenta	-92.455336	240.370377	First	Female	GSE131945	Pure	0.9909429
GSM2674512_10005833024_R05C01	Placenta	-108.39993	230.349804	Term	Female	GSE100197	Pure	0.99120719
GSM3396868_9992571127_R02C01	Placenta	-57.189294	243.695653	Term	Female	GSE120250	Pure	0.99150473
GSM3396857_9992576033_R04C02	Placenta	-25.800654	234.390365	Term	Male	GSE120250	Pure	0.99167877
GSM1931597_9296930103_R01C01	Placenta	-68.578288	240.779813	Term	Male	GSE74738	Pure	0.99168563
GSM1617000_8795207029_R02C02	Placenta	-23.3982	232.411924	First	Male	GSE66210	Pure	0.99172636
GSM1947110_5806636086_R02C01	Placenta	17.2868032	174.726041	Term	Male	GSE75248	Pure	0.99193574
GSM3396845_9992576033_R03C01	Placenta	-15.818507	223.973453	Term	Female	GSE120250	Pure	0.99195319
GSM2674509_9977525013_R03C02	Placenta	-45.877543	233.984717	Term	Male	GSE100197	Pure	0.99229128

GSM1931599_9285451020_R01C01	Placenta	-77.197788	231.421399	Term	Female	GSE74738	Pure	0.99242458
GSM3396830_9992571114_R05C02	Placenta	-10.038681	209.738536	Term	Female	GSE120250	Pure	0.99253315
GSM1702195_9296930114_R02C01	Placenta	-117.33048	212.423786	Second	Female	GSE69502	Pure	0.99265744
GSM3396867_9992571114_R06C02	Placenta	9.30201735	175.425394	Term	Female	GSE120250	Pure	0.99278733
GSM1931601_9266441046_R04C01	Placenta	-50.945113	226.44507	Term	Male	GSE74738	Pure	0.99295481
GSM3396861_9992576207_R01C02	Placenta	10.0004896	169.390698	Term	Female	GSE120250	Pure	0.99299868
GSM1947287_5806417045_R04C02	Placenta	-88.526495	221.758206	Term	Male	GSE75248	Pure	0.99301975
GSM3396866_9992576159_R04C01	Placenta	-7.6527454	196.061088	Term	Male	GSE120250	Pure	0.99316222
GSM3396832_9992576033_R03C02	Placenta	-15.343076	198.750998	Term	Female	GSE120250	Pure	0.99349971
GSM3396842_9992576160_R01C02	Placenta	8.06936571	161.406762	Term	Female	GSE120250	Pure	0.99353349
GSM2674507_9977525015_R05C01	Placenta	-115.89682	202.804202	Term	Female	GSE100197	Pure	0.99365086
GSM1843084_3999984151_R01C01	Placenta	32.5705765	-307.09618	Term	Male	GSE71678	Pure	0.99369313
GSM1947043_5806636061_R04C01	Placenta	-26.40977	205.000177	Term	Female	GSE75248	Pure	0.99369654
GSM3179718_200889820007_R07C01	Placenta	-158.33975	169.094777	Term	Male	GSE115508	Pure	0.99411828
GSM1616998_8795207029_R02C01	Placenta	18.2364521	115.779295	First	Male	GSE66210	Pure	0.99412303
GSM1842925_3999997014_R05C01	Placenta	20.5073473	100.532229	Term	Female	GSE71678	Pure	0.9943032
GSM3396833_9992571130_R03C02	Placenta	-19.26013	185.289664	Term	Male	GSE120250	Pure	0.99438902
GSM1947073_5806636039_R05C01	Placenta	35.1738069	-14.05639	Term	Male	GSE75248	Pure	0.99443174
GSM1947160_5806636035_R01C01	Placenta	-13.81815	175.895971	Term	Female	GSE75248	Pure	0.9945241
GSM1702185_7973201026_R05C02	Placenta	36.3389104	-169.00867	Second	Female	GSE69502	Pure	0.99456785
GSM1931585_9296930098_R03C01	Placenta	-105.60111	196.074299	Term	Male	GSE74738	Pure	0.99457213
GSM3396853_9992576033_R01C01	Placenta	-34.549547	192.796574	Term	Male	GSE120250	Pure	0.99464977
GSM1947269_5806636051_R02C02	Placenta	-15.025549	169.840542	Term	Male	GSE75248	Pure	0.99484351
GSM1931591_9266441046_R01C01	Placenta	-132.21094	178.592324	Term	Female	GSE74738	Pure	0.99485022
GSM1947223_5806636042_R06C01	Placenta	17.6593805	89.3929792	Term	Female	GSE75248	Pure	0.99488811
GSM3396848_9992571130_R05C02	Placenta	-24.258306	178.209278	Term	Male	GSE120250	Pure	0.9949378
GSM1702149_7970368050_R06C02	Placenta	33.6268885	-187.73662	Second	Male	GSE69502	Pure	0.99498129

GSM1947118_5806636090_R03C01	Placenta	-10.994389	159.664223	Term	Female	GSE75248	Pure	0.99501435
GSM1702147_7970368050_R05C01	Placenta	32.0788728	-218.12213	Second	Male	GSE69502	Pure	0.99505367
GSM1947102_5806636039_R05C02	Placenta	21.1226677	65.1033429	Term	Male	GSE75248	Pure	0.99505527
GSM2674514_7970368023_R05C02	Placenta	31.6874602	-217.22472	Term	Female	GSE100197	Pure	0.99513511
GSM2674505_9296930098_R03C01	Placenta	-99.177119	189.441557	Term	Male	GSE100197	Pure	0.99516756
GSM3396844_9992571127_R04C01	Placenta	-63.846059	190.961573	Term	Male	GSE120250	Pure	0.99529627
GSM2674508_9266441156_R03C01	Placenta	-108.10529	184.143692	Term	Female	GSE100197	Pure	0.99529792
GSM1947261_5806636063_R03C01	Placenta	-18.300688	161.702874	Term	Female	GSE75248	Pure	0.99532411
GSM1947265_5806636083_R02C02	Placenta	-13.241971	153.921073	Term	Male	GSE75248	Pure	0.99534567
GSM3396851_9992576207_R06C02	Placenta	20.6070265	52.0404348	Term	Male	GSE120250	Pure	0.99535498
GSM1947195_6008581005_R01C01	Placenta	28.5257254	-25.016547	Term	Female	GSE75248	Pure	0.99547206
GSM1947158_5806636036_R05C02	Placenta	-26.117781	163.539752	Term	Female	GSE75248	Pure	0.99559767
GSM1947209_5806417009_R06C02	Placenta	13.8051006	70.6084644	Term	Male	GSE75248	Pure	0.99565862
GSM1947215_5806636034_R03C01	Placenta	-42.715999	173.315869	Term	Male	GSE75248	Pure	0.99575168
GSM1947034_5806636041_R05C01	Placenta	20.004645	23.2797129	Term	Female	GSE75248	Pure	0.99586612
GSM1947080_5859594028_R06C01	Placenta	15.2820354	51.243238	Term	Male	GSE75248	Pure	0.99588027
GSM1931570_9296930123_R06C01	Placenta	-46.4387	171.860279	Term	Male	GSE74738	Pure	0.99589545
GSM3396847_9992571127_R01C02	Placenta	-13.828246	130.636023	Term	Male	GSE120250	Pure	0.99609909
GSM2674501_9285451020_R05C01	Placenta	-77.657014	174.136893	Term	Female	GSE100197	Pure	0.99614147
GSM1947142_5806636063_R06C02	Placenta	-19.863368	137.652445	Term	Male	GSE75248	Pure	0.99618069
GSM3396841_9992576207_R04C02	Placenta	13.2894545	42.3932937	Term	Male	GSE120250	Pure	0.99619336
GSM2674502_9285451020_R01C02	Placenta	-87.092094	172.394111	Term	Male	GSE100197	Pure	0.99620039
GSM1617006_8795194156_R02C01	Placenta	-75.946408	172.080004	First	Female	GSE66210	Pure	0.99622709
GSM2674497_9296930123_R06C01	Placenta	-53.887458	165.800327	Term	Male	GSE100197	Pure	0.9962678
GSM1931596_9266441046_R02C01	Placenta	-84.951697	170.306766	Term	Male	GSE74738	Pure	0.99629981
GSM2674513_9266441046_R02C01	Placenta	-77.937738	168.983663	Term	Male	GSE100197	Pure	0.99635937
GSM1931576_9285451020_R05C01	Placenta	-83.682402	166.789795	Term	Female	GSE74738	Pure	0.99644972

GSM1931571_9266441156_R05C02	Placenta	-65.781436	164.622978	Term	Female	GSE74738	Pure	0.99646314
GSM1617004_7668610068_R05C01	Placenta	-48.363772	156.991184	First	Female	GSE66210	Pure	0.99647395
GSM1947194_5806636034_R01C01	Placenta	-54.219651	156.737481	Term	Male	GSE75248	Pure	0.99659035
GSM2674498_9266441156_R05C02	Placenta	-64.898025	159.792148	Term	Female	GSE100197	Pure	0.99662857
GSM1947093_5806636038_R03C02	Placenta	12.0277985	16.7590738	Term	Male	GSE75248	Pure	0.99664573
GSM1616999_8795207029_R05C01	Placenta	1.77291787	61.0924739	First	Male	GSE66210	Pure	0.99672172
GSM1947141_5806636090_R03C02	Placenta	-26.950095	124.570256	Term	Male	GSE75248	Pure	0.99680494
GSM1947271_5806636061_R05C02	Placenta	-37.656694	136.826626	Term	Female	GSE75248	Pure	0.99682151
GSM1842930_3999997014_R04C02	Placenta	-114.75207	150.660958	Term	Female	GSE71678	Pure	0.99685778
GSM1947246_5806636062_R05C01	Placenta	-44.001725	140.662611	Term	Male	GSE75248	Pure	0.99687554
GSM1931569_9296930098_R01C02	Placenta	-53.745624	147.399057	Term	Female	GSE74738	Pure	0.99687616
GSM2674504_9285451059_R04C01	Placenta	-108.76218	151.159304	Term	Female	GSE100197	Pure	0.9969121
GSM1931577_9285451020_R01C02	Placenta	-86.374659	154.114654	Term	Male	GSE74738	Pure	0.99692758
GSM1842895_3999997150_R04C02	Placenta	5.15561854	26.8482101	Term	Male	GSE71678	Pure	0.99700198
GSM1617008_8795194156_R05C01	Placenta	-119.17898	144.520642	First	Female	GSE66210	Pure	0.99703784
GSM1947163_5806636026_R03C02	Placenta	-64.48262	144.580432	Term	Female	GSE75248	Pure	0.99710963
GSM1947126_5859594028_R02C02	Placenta	12.4464651	-39.336048	Term	Female	GSE75248	Pure	0.99715554
GSM1931573_9296930098_R01C01	Placenta	-101.77108	146.081592	Term	Male	GSE74738	Pure	0.99715605
GSM1947044_5806636042_R04C01	Placenta	-6.7988223	61.4840798	Term	Female	GSE75248	Pure	0.99718964
GSM2674496_9296930098_R01C02	Placenta	-74.57077	142.867771	Term	Female	GSE100197	Pure	0.99724852
GSM1947146_5806417045_R06C02	Placenta	-89.514452	140.929154	Term	Male	GSE75248	Pure	0.99735573
GSM1947092_6008581008_R03C01	Placenta	9.8038296	-41.92739	Term	Female	GSE75248	Pure	0.99736159
GSM1947117_6008581008_R03C02	Placenta	14.1332216	-105.51937	Term	Male	GSE75248	Pure	0.99738082
GSM1947039_5806636086_R02C02	Placenta	-8.1276235	35.215001	Term	Male	GSE75248	Pure	0.99761074
GSM1842879_3999984102_R06C01	Placenta	-8.8094318	31.7659205	Term	Female	GSE71678	Pure	0.99768183
GSM2674506_9285451059_R02C01	Placenta	-116.80055	125.272994	Term	Male	GSE100197	Pure	0.99770586
GSM1947089_5806636039_R06C02	Placenta	1.86838591	-27.758848	Term	Male	GSE75248	Pure	0.99773593

GSM1947029_5806417045_R01C02	Placenta	-26.612002	75.7912803	Term	Male	GSE75248	Pure	0.9977462
GSM2674503_9266441156_R06C02	Placenta	-97.526781	126.627436	Term	Male	GSE100197	Pure	0.99775203
GSM1947233_5859594025_R01C01	Placenta	-74.380357	123.227058	Term	Female	GSE75248	Pure	0.99775562
GSM1947078_6008581005_R03C02	Placenta	10.846538	-206.86703	Term	Male	GSE75248	Pure	0.99775709
GSM1702135_7970368112_R06C01	Placenta	2.92646641	-383.80465	Second	Female	GSE69502	Pure	0.99775789
GSM1702082_7970368015_R03C01	Placenta	-9.3717623	26.4981248	Second	Female	GSE69502	Pure	0.99776601
GSM2674500_9296930103_R06C02	Placenta	-96.965599	124.815919	Term	Female	GSE100197	Pure	0.99779874
GSM1947210_5859594028_R02C01	Placenta	-1.7131177	-14.218855	Term	Male	GSE75248	Pure	0.99781119
GSM3396863_9992576022_R06C01	Placenta	-26.550657	70.2672435	Term	Male	GSE120250	Pure	0.99782674
GSM1931581_9285451059_R04C01	Placenta	-120.4583	118.190124	Term	Female	GSE74738	Pure	0.99787429
GSM3396849_9992576207_R02C01	Placenta	-26.641367	58.8178097	Term	Female	GSE120250	Pure	0.99798723
GSM1842851_3999984095_R02C01	Placenta	-91.602156	114.678014	Term	Male	GSE71678	Pure	0.99803049
GSM1947040_5859594006_R01C02	Placenta	-56.299989	98.4561661	Term	Female	GSE75248	Pure	0.99803899
GSM1947151_5859720025_R05C02	Placenta	-9.8599175	-1.2115703	Term	Female	GSE75248	Pure	0.99805939
GSM1947238_5806636062_R01C02	Placenta	-39.426855	75.6761856	Term	Female	GSE75248	Pure	0.99808368
GSM1947278_6190781088_R04C02	Placenta	5.84602702	-176.06525	Term	Male	GSE75248	Pure	0.99809954
GSM1947216_5806636042_R01C02	Placenta	-35.705275	63.348291	Term	Male	GSE75248	Pure	0.99816651
GSM1931575_9296930103_R06C02	Placenta	-100.3875	108.75218	Term	Female	GSE74738	Pure	0.99816707
GSM3396870_9992571114_R05C01	Placenta	-24.700273	38.2144406	Term	Female	GSE120250	Pure	0.99817727
GSM3396856_9992576160_R02C01	Placenta	-44.406906	75.8908001	Term	Male	GSE120250	Pure	0.99818448
201332340153_R05C01	Placenta	-89.910558	104.631384	First	Female	GSE131945	Pure	0.99822929
GSM1843130_3999997049_R01C01	Placenta	-10.056756	-32.15008	Term	Male	GSE71678	Pure	0.99830409
GSM1617001_8795207029_R05C02	Placenta	-62.912017	85.732287	First	Male	GSE66210	Pure	0.99833277
GSM1947149_5806636085_R04C02	Placenta	-33.782738	44.2209413	Term	Male	GSE75248	Pure	0.99834883
GSM1947200_6008581008_R05C01	Placenta	0.67247973	-152.49253	Term	Male	GSE75248	Pure	0.99835143
GSM1947024_5806417009_R01C01	Placenta	-4.5408893	-81.573361	Term	Female	GSE75248	Pure	0.99836103
GSM1947123_5806417076_R05C01	Placenta	-57.178177	74.8536525	Term	Male	GSE75248	Pure	0.99841124

GSM1947267_5806417081_R02C02	Placenta	-63.528877	80.5950259	Term	Female	GSE75248	Pure	0.99841452
GSM1947075_5859594028_R06C02	Placenta	-12.353375	-39.044162	Term	Female	GSE75248	Pure	0.99843113
GSM1947047_5859594009_R02C02	Placenta	-27.198384	14.7643154	Term	Female	GSE75248	Pure	0.99847516
GSM1944962_9376561070_R02C02	Placenta	-129.02721	88.6660091	Term	Male	GSE75196	Pure	0.99848902
GSM1947086_5806636042_R03C01	Placenta	-59.2048	70.3949598	Term	Male	GSE75248	Pure	0.99849682
GSM1947082_5806417045_R03C02	Placenta	-45.900884	48.7820316	Term	Male	GSE75248	Pure	0.99854542
GSM1947229_5806417076_R03C02	Placenta	-47.939434	49.3513985	Term	Male	GSE75248	Pure	0.99857407
GSM2589567_10005833024_R06C02	Placenta	-34.206227	22.1265027	Term	Male	GSE98224	Pure	0.99857611
GSM1702102_7970368054_R05C01	Placenta	-9.2416168	-389.39369	Second	Male	GSE69502	Pure	0.99863699
GSM1947354_3999997077_R02C01	Placenta	-18.628999	-51.010298	Term	Male	GSE75248	Pure	0.99869651
GSM2589570_10005833038_R06C01	Placenta	-34.740357	2.5614065	Term	Male	GSE98224	Pure	0.99874615
GSM1947247_5806417077_R03C02	Placenta	-28.749368	-16.126462	Term	Male	GSE75248	Pure	0.99874724
GSM1947036_5806636029_R01C02	Placenta	-21.820937	-44.358597	Term	Female	GSE75248	Pure	0.9987475
GSM1947111_6190781088_R04C01	Placenta	-28.73004	-21.155584	Term	Male	GSE75248	Pure	0.99877913
GSM1947132_5806636019_R03C01	Placenta	-36.500394	2.66301313	Term	Female	GSE75248	Pure	0.99878049
201332340172_R03C01	Placenta	-91.220234	67.7108951	First	Male	GSE131945	Pure	0.99879957
GSM1842835_3999984083_R03C02	Placenta	-282.13588	-45.502269	Term	Male	GSE71678	Pure	0.99881362
GSM1947097_5806417077_R02C01	Placenta	-76.842318	57.4148496	Term	Male	GSE75248	Pure	0.99882885
GSM1947016_3999997080_R04C02	Placenta	-15.219953	-106.97322	Term	Male	GSE75248	Pure	0.99883692
GSM1947091_5859594009_R05C01	Placenta	-48.965351	21.9195512	Term	Female	GSE75248	Pure	0.99884572
GSM1842945_3999997033_R01C01	Placenta	-296.02227	-67.668351	Term	Male	GSE71678	Pure	0.99894601
GSM1944953_9376561054_R05C02	Placenta	-49.344781	8.08602307	Term	Male	GSE75196	Pure	0.99895622
GSM1947259_5806417009_R02C02	Placenta	-27.962931	-60.457224	Term	Male	GSE75248	Pure	0.99897272
GSM2589572_10005833110_R05C02	Placenta	-40.930569	-20.584768	Term	Female	GSE98224	Pure	0.99901096
GSM1702209_9296930117_R06C02	Placenta	-30.09601	-64.657988	Second	Female	GSE69502	Pure	0.99903514
GSM1617009_8795194156_R05C02	Placenta	-74.945059	33.8599813	First	Female	GSE66210	Pure	0.99903702
GSM1947245_5859594006_R05C01	Placenta	-26.967834	-86.1214	Term	Male	GSE75248	Pure	0.99905627

GSM1842908_3999997011_R06C02	Placenta	-64.98652	19.4851351	Term	Female	GSE71678	Pure	0.999061
GSM1947095_6008581003_R04C01	Placenta	-21.299221	-129.4302	Term	Male	GSE75248	Pure	0.99906465
GSM1947022_5806417081_R02C01	Placenta	-44.463923	-21.794502	Term	Female	GSE75248	Pure	0.99907291
GSM1947293_3999997076_R06C01	Placenta	-18.828776	-159.95826	Term	Male	GSE75248	Pure	0.9990731
GSM1947026_5806636019_R06C02	Placenta	-42.503578	-40.072955	Term	Female	GSE75248	Pure	0.99913799
GSM1947154_5806636020_R02C01	Placenta	-79.799247	25.3996293	Term	Male	GSE75248	Pure	0.9991395
GSM1947370_3999997077_R01C01	Placenta	-34.308284	-71.406961	Term	Male	GSE75248	Pure	0.9991413
GSM1947032_5806417077_R05C02	Placenta	-45.272739	-35.457604	Term	Female	GSE75248	Pure	0.9991558
GSM1947114_5806636030_R03C02	Placenta	-74.201446	16.8086264	Term	Female	GSE75248	Pure	0.99916105
GSM1947279_5806417023_R04C02	Placenta	-101.1402	35.9704361	Term	Male	GSE75248	Pure	0.99916159
GSM1947262_5806417064_R03C02	Placenta	-72.169674	13.9411612	Term	Male	GSE75248	Pure	0.99916467
GSM1947218_5806417045_R03C01	Placenta	-63.135332	-2.5034044	Term	Female	GSE75248	Pure	0.99918864
GSM1947148_5806636030_R04C01	Placenta	-89.373524	24.7542297	Term	Female	GSE75248	Pure	0.99920089
GSM1947256_5859720011_R04C01	Placenta	-35.429368	-87.483117	Term	Male	GSE75248	Pure	0.99921861
GSM1842845_3999984084_R01C02	Placenta	-332.32956	-128.24221	Term	Male	GSE71678	Pure	0.99924228
GSM1944960_9376561070_R06C01	Placenta	-116.96869	29.4727678	Term	Female	GSE75196	Pure	0.9992508
GSM1842924_3999997014_R04C01	Placenta	-273.85477	-62.164838	Term	Female	GSE71678	Pure	0.99925141
GSM1842914_3999997013_R06C01	Placenta	-174.82024	17.126789	Term	Male	GSE71678	Pure	0.99925395
GSM1947404_3999997076_R04C01	Placenta	-61.246086	-19.660486	Term	Male	GSE75248	Pure	0.99926484
GSM1947361_3999997074_R05C02	Placenta	-34.804607	-111.39564	Term	Male	GSE75248	Pure	0.99928152
GSM1944959_9376561070_R05C01	Placenta	-65.207853	-16.38998	Term	Female	GSE75196	Pure	0.99928498
GSM1947225_5859720012_R01C01	Placenta	-50.175771	-54.261254	Term	Female	GSE75248	Pure	0.99930205
GSM1947145_5806636030_R05C01	Placenta	-71.559913	-10.332306	Term	Female	GSE75248	Pure	0.99930595
GSM1947014_3999997072_R03C02	Placenta	-60.313392	-31.721011	Term	Female	GSE75248	Pure	0.99931403
GSM1947234_6008581009_R02C01	Placenta	-35.997284	-128.12906	Term	Female	GSE75248	Pure	0.999343
GSM1842970_3999997039_R04C01	Placenta	-152.8403	15.1728411	Term	Male	GSE71678	Pure	0.99934876
GSM1947124_6008581008_R01C01	Placenta	-39.366537	-111.6885	Term	Female	GSE75248	Pure	0.99935102

GSM1947133_6008581008_R04C01	Placenta	-34.09295	-149.4033	Term	Female	GSE75248	Pure	0.99935992
GSM1947115_5806636030_R02C02	Placenta	-96.906419	5.05903642	Term	Female	GSE75248	Pure	0.99936525
GSM1931582_7973201026_R03C02	Placenta	-82.075453	-8.5399709	Term	Female	GSE74738	Pure	0.99936772
GSM1947369_3999997072_R02C02	Placenta	-72.111626	-23.891521	Term	Female	GSE75248	Pure	0.99937742
GSM1947272_5806417023_R01C01	Placenta	-99.622433	4.38975768	Term	Female	GSE75248	Pure	0.99937948
GSM1947224_5806417064_R02C01	Placenta	-48.510069	-85.189305	Term	Male	GSE75248	Pure	0.99939128
GSM1944963_9376561070_R03C02	Placenta	-49.581545	-85.914069	Term	Male	GSE75196	Pure	0.99940557
GSM1947401_3999997093_R02C02	Placenta	-79.256169	-21.336158	Term	Male	GSE75248	Pure	0.99941357
GSM1947352_3999997091_R02C02	Placenta	-113.95434	2.86316343	Term	Male	GSE75248	Pure	0.99942888
GSM1947317_3999997100_R04C01	Placenta	-52.745297	-83.16376	Term	Female	GSE75248	Pure	0.99943101
GSM1947113_3999997076_R01C01	Placenta	-162.68419	-2.2204406	Term	Female	GSE75248	Pure	0.99945848
GSM1947231_5806417023_R02C02	Placenta	-58.856002	-78.048223	Term	Male	GSE75248	Pure	0.99947372
GSM1947284_5806417064_R01C02	Placenta	-66.501271	-58.600108	Term	Female	GSE75248	Pure	0.99947395
GSM1702208_9296930117_R06C01	Placenta	-134.90368	-1.5819437	Second	Female	GSE69502	Pure	0.99947879
GSM1842843_3999984084_R05C01	Placenta	-263.1771	-73.41027	Term	Male	GSE71678	Pure	0.99948656
GSM3396840_9992576022_R04C01	Placenta	-82.213296	-34.618663	Term	Male	GSE120250	Pure	0.99948782
GSM1947311_3999997076_R01C02	Placenta	-74.400367	-48.60846	Term	Male	GSE75248	Pure	0.9994943
GSM1947143_5806636030_R01C02	Placenta	-82.171678	-36.532071	Term	Female	GSE75248	Pure	0.99949516
GSM1947038_6008581014_R04C02	Placenta	-38.870638	-191.41373	Term	Female	GSE75248	Pure	0.99949711
GSM1947409_3999997093_R01C02	Placenta	-102.95424	-16.173854	Term	Female	GSE75248	Pure	0.99949828
GSM1947030_5806417076_R02C01	Placenta	-145.29701	-5.3138704	Term	Male	GSE75248	Pure	0.99949969
GSM1947384_3999997072_R04C01	Placenta	-107.43008	-15.708689	Term	Male	GSE75248	Pure	0.99950955
GSM1947324_3999997074_R05C01	Placenta	-62.908024	-79.994117	Term	Female	GSE75248	Pure	0.99951333
GSM1947286_5806417009_R02C01	Placenta	-79.30989	-46.401229	Term	Female	GSE75248	Pure	0.99951616
GSM1947257_5806636019_R04C01	Placenta	-78.167299	-48.386525	Term	Male	GSE75248	Pure	0.99951663
GSM1842932_3999997014_R06C02	Placenta	-106.90201	-17.650582	Term	Male	GSE71678	Pure	0.99951701
GSM1947211_5859594006_R06C02	Placenta	-80.099824	-46.41856	Term	Female	GSE75248	Pure	0.99952069

GSM2589576_9977525013_R06C01	Placenta	-49.089155	-142.34539	Term	Male	GSE98224	Pure	0.99953629
GSM1702206_9296930117_R05C01	Placenta	-66.092901	-79.636656	Second	Female	GSE69502	Pure	0.99953647
GSM1947134_5859720025_R04C02	Placenta	-101.50132	-26.176666	Term	Male	GSE75248	Pure	0.99953744
GSM3396834_9992576207_R03C01	Placenta	-113.0926	-18.809363	Term	Female	GSE120250	Pure	0.99953808
GSM1947270_5806417077_R05C01	Placenta	-94.357032	-33.235421	Term	Male	GSE75248	Pure	0.99953966
GSM1944955_9376561070_R01C01	Placenta	-83.391566	-50.538303	Term	Female	GSE75196	Pure	0.99955263
GSM1947408_3999997077_R06C01	Placenta	-54.449945	-126.12199	Term	Female	GSE75248	Pure	0.99955525
GSM1842892_3999997150_R01C02	Placenta	-72.256129	-72.687059	Term	Female	GSE71678	Pure	0.99955847
GSM1947398_3999997076_R03C01	Placenta	-127.03793	-21.623739	Term	Male	GSE75248	Pure	0.99957485
GSM1947327_3999997091_R04C01	Placenta	-84.371256	-56.991716	Term	Female	GSE75248	Pure	0.99957863
GSM1702121_7970368097_R03C01	Placenta	-98.703499	-39.846769	Second	Female	GSE69502	Pure	0.99958076
GSM1947345_3999997051_R06C02	Placenta	-94.702102	-44.503752	Term	Female	GSE75248	Pure	0.99958243
GSM1842844_3999984084_R06C01	Placenta	-173.72978	-26.154287	Term	Male	GSE71678	Pure	0.99958293
GSM1944956_9376561070_R02C01	Placenta	-105.43789	-37.914284	Term	Female	GSE75196	Pure	0.99959536
GSM1944958_9376561070_R04C01	Placenta	-86.96105	-60.141163	Term	Female	GSE75196	Pure	0.99960055
GSM1947383_3999997074_R02C01	Placenta	-67.897183	-100.96153	Term	Male	GSE75248	Pure	0.99960353
GSM1947334_3999997076_R05C02	Placenta	-72.528059	-91.552041	Term	Female	GSE75248	Pure	0.99961027
GSM1944954_9376561054_R06C02	Placenta	-84.165782	-73.097774	Term	Female	GSE75196	Pure	0.99962475
GSM1947393_3999997077_R05C02	Placenta	-73.455231	-95.676159	Term	Male	GSE75248	Pure	0.99962552
GSM1842962_3999997037_R01C02	Placenta	-309.7527	-139.48248	Term	Female	GSE71678	Pure	0.99962683
GSM1947331_3999997074_R06C02	Placenta	-77.459391	-90.130235	Term	Male	GSE75248	Pure	0.99963458
GSM1947291_3999997074_R01C02	Placenta	-105.1884	-51.614187	Term	Male	GSE75248	Pure	0.9996409
GSM1842868_3999984096_R01C02	Placenta	-269.73391	-102.35988	Term	Female	GSE71678	Pure	0.99964938
GSM1944961_9376561070_R01C02	Placenta	-152.09784	-38.186664	Term	Male	GSE75196	Pure	0.99965707
GSM1947335_3999997072_R01C01	Placenta	-79.437902	-100.34612	Term	Female	GSE75248	Pure	0.99966775
GSM1947395_3999997077_R04C01	Placenta	-95.018265	-72.968436	Term	Male	GSE75248	Pure	0.99966886
GSM1947077_6008581005_R06C01	Placenta	-85.762168	-88.514966	Term	Male	GSE75248	Pure	0.99967045

GSM1931574_7970368100_R02C02	Placenta	-74.42224	-114.23888	Term	Male	GSE74738	Pure	0.99967051
GSM1947350_3999997110_R01C02	Placenta	-63.180411	-153.53209	Term	Female	GSE75248	Pure	0.99967128
GSM1931578_9296930098_R06C01	Placenta	-125.27224	-48.036866	Term	Female	GSE74738	Pure	0.99967274
GSM1947388_3999997074_R04C02	Placenta	-121.78786	-54.5327	Term	Male	GSE75248	Pure	0.99968732
GSM1947330_3999997074_R01C01	Placenta	-102.15669	-72.140281	Term	Male	GSE75248	Pure	0.99968997
GSM1944951_9376561054_R03C02	Placenta	-112.58497	-62.297829	Term	Female	GSE75196	Pure	0.99969124
GSM1947326_3999997100_R04C02	Placenta	-139.11964	-50.173325	Term	Male	GSE75248	Pure	0.99969511
GSM1947230_6008581026_R02C02	Placenta	-127.42987	-59.258839	Term	Male	GSE75248	Pure	0.99970987
GSM1944964_9376561070_R04C02	Placenta	-111.78816	-72.292164	Term	Male	GSE75196	Pure	0.99971553
GSM1944952_9376561054_R04C02	Placenta	-103.65656	-81.267086	Term	Male	GSE75196	Pure	0.99971633
GSM1947290_3999997091_R03C01	Placenta	-86.674977	-115.18968	Term	Male	GSE75248	Pure	0.99972845
GSM1947254_5806417077_R01C01	Placenta	-71.082986	-163.03759	Term	Female	GSE75248	Pure	0.99973123
GSM1947296_3999997076_R02C02	Placenta	-82.577787	-133.46753	Term	Female	GSE75248	Pure	0.99974244
GSM1947349_3999997110_R06C02	Placenta	-99.224178	-101.06713	Term	Male	GSE75248	Pure	0.99974513
GSM1944957_9376561070_R03C01	Placenta	-112.5934	-86.343083	Term	Female	GSE75196	Pure	0.99974959
GSM1947139_5806417081_R06C02	Placenta	-96.757887	-107.81318	Term	Female	GSE75248	Pure	0.99975024
GSM1947379_3999997072_R01C02	Placenta	-95.329398	-114.11072	Term	Male	GSE75248	Pure	0.99975689
GSM1842850_3999984095_R01C01	Placenta	-293.31952	-148.65064	Term	Female	GSE71678	Pure	0.99976509
GSM1947312_3999997091_R05C02	Placenta	-97.430298	-116.0511	Term	Female	GSE75248	Pure	0.99976658
GSM1947159_6008581014_R05C01	Placenta	-69.835721	-208.81945	Term	Female	GSE75248	Pure	0.99977463
GSM1944965_9376561070_R05C02	Placenta	-111.83586	-102.66069	Term	Male	GSE75196	Pure	0.99977986
GSM1947292_3999997100_R01C01	Placenta	-109.41577	-109.6872	Term	Male	GSE75248	Pure	0.99978681
GSM1947329_3999997076_R03C02	Placenta	-135.01452	-89.014039	Term	Male	GSE75248	Pure	0.99979086
GSM1947366_3999997072_R03C01	Placenta	-138.27675	-89.119089	Term	Female	GSE75248	Pure	0.99979444
GSM1842979_3999997059_R01C01	Placenta	-116.12246	-107.29582	Term	Male	GSE71678	Pure	0.99979651
GSM1947364_3999997074_R02C02	Placenta	-144.97081	-90.174302	Term	Female	GSE75248	Pure	0.99980226
GSM1842904_3999997011_R02C02	Placenta	-195.58017	-97.062985	Term	Male	GSE71678	Pure	0.99981603

GSM1947193_6008581026_R05C02	Placenta	-128.77199	-110.30101	Term	Male	GSE75248	Pure	0.99982149
GSM1843115_3999997017_R03C02	Placenta	-115.64265	-124.18075	Term	Female	GSE71678	Pure	0.9998218
GSM1947356_3999997074_R03C02	Placenta	-132.09248	-112.25027	Term	Female	GSE75248	Pure	0.99982872
GSM1843017_3999984029_R05C01	Placenta	-94.710512	-169.29305	Term	Male	GSE71678	Pure	0.99982901
GSM1843123_3999997018_R05C01	Placenta	-145.10532	-105.56421	Term	Male	GSE71678	Pure	0.99983064
GSM1947239_5806636037_R03C02	Placenta	-89.0198	-188.23147	Term	Male	GSE75248	Pure	0.99983095
GSM1947392_3999997114_R03C02	Placenta	-100.30464	-163.35993	Term	Male	GSE75248	Pure	0.9998372
GSM1947355_3999997091_R06C02	Placenta	-145.41331	-112.1494	Term	Female	GSE75248	Pure	0.99984149
GSM1843139_3999997049_R05C02	Placenta	-228.33327	-122.64068	Term	Male	GSE71678	Pure	0.99984162
GSM1947333_3999997077_R04C02	Placenta	-125.60632	-127.05557	Term	Female	GSE75248	Pure	0.99984168
GSM1947397_3999997093_R05C02	Placenta	-174.19575	-105.89375	Term	Male	GSE75248	Pure	0.99984174
GSM1842995_3999984004_R06C01	Placenta	-94.233083	-183.80345	Term	Female	GSE71678	Pure	0.99984175
GSM1947328_3999997080_R05C01	Placenta	-113.06974	-143.98935	Term	Male	GSE75248	Pure	0.99984269
GSM1947381_3999997110_R01C01	Placenta	-121.29757	-132.86681	Term	Female	GSE75248	Pure	0.99984296
GSM1947342_3999997110_R05C01	Placenta	-97.254195	-177.26142	Term	Male	GSE75248	Pure	0.99984351
GSM1947128_6008581028_R02C02	Placenta	-92.046518	-197.32914	Term	Female	GSE75248	Pure	0.9998475
GSM1947353_3999997080_R04C01	Placenta	-106.95866	-165.96407	Term	Female	GSE75248	Pure	0.99985451
GSM1947405_3999997093_R02C01	Placenta	-116.45998	-150.54854	Term	Male	GSE75248	Pure	0.99985602
GSM1947338_3999997072_R06C01	Placenta	-96.865741	-196.30743	Term	Female	GSE75248	Pure	0.99985905
GSM1944966_9376561070_R06C02	Placenta	-177.98252	-116.61168	Term	Female	GSE75196	Pure	0.99985998
GSM1947357_3999997091_R06C01	Placenta	-147.26154	-133.6099	Term	Female	GSE75248	Pure	0.99987238
GSM1947358_3999997100_R06C02	Placenta	-138.04764	-145.91473	Term	Male	GSE75248	Pure	0.99987863
GSM1947304_3999997072_R05C02	Placenta	-150.08926	-140.78248	Term	Female	GSE75248	Pure	0.99988268
GSM1947348_3999997110_R04C01	Placenta	-120.22039	-173.4286	Term	Male	GSE75248	Pure	0.9998838
GSM1931568_9285451059_R06C02	Placenta	-131.16089	-167.93345	Term	Male	GSE74738	Pure	0.99989269
GSM1843136_3999997049_R02C02	Placenta	-251.73496	-166.48352	Term	Male	GSE71678	Pure	0.99989675
GSM1842985_3999997059_R01C02	Placenta	-328.87432	-232.04558	Term	Female	GSE71678	Pure	0.99989704

GSM1947303_3999997093_R06C01	Placenta	-108.8205	-217.30496	Term	Female	GSE75248	Pure	0.99989776
GSM1947332_3999997114_R01C02	Placenta	-116.5351	-199.1614	Term	Female	GSE75248	Pure	0.99989811
GSM2674499_7970368050_R03C02	Placenta	-122.61268	-191.02666	Term	Male	GSE100197	Pure	0.99990068
GSM1843041_3999984049_R05C01	Placenta	-126.28687	-185.98879	Term	Female	GSE71678	Pure	0.99990151
GSM1842853_3999984095_R04C01	Placenta	-294.02187	-204.1751	Term	Female	GSE71678	Pure	0.99990558
GSM1842951_3999997033_R01C02	Placenta	-229.78189	-164.13145	Term	Male	GSE71678	Pure	0.99990951
GSM1947347_3999997074_R03C01	Placenta	-175.48734	-156.44241	Term	Female	GSE75248	Pure	0.99990952
GSM1947313_3999997093_R05C01	Placenta	-108.34218	-246.45808	Term	Male	GSE75248	Pure	0.99991253
GSM1843049_3999984068_R01C01	Placenta	-134.67386	-189.94664	Term	Male	GSE71678	Pure	0.99991307
GSM1947399_3999997080_R02C01	Placenta	-169.78747	-166.27872	Term	Male	GSE75248	Pure	0.99991682
GSM1947322_3999997110_R06C01	Placenta	-145.12795	-184.61836	Term	Male	GSE75248	Pure	0.99991802
201414140063_R07C01	Placenta	-309.89773	-229.37247	First	Female	GSE131945	Pure	0.99992124
GSM1843117_3999997017_R05C02	Placenta	-215.58855	-176.39417	Term	Male	GSE71678	Pure	0.99992739
GSM1947406_3999997080_R03C01	Placenta	-190.14136	-177.9324	Term	Female	GSE75248	Pure	0.99993017
GSM1842980_3999997059_R02C01	Placenta	-189.48282	-179.37457	Term	Male	GSE71678	Pure	0.99993121
GSM1947387_3999997114_R06C02	Placenta	-123.79423	-253.36029	Term	Male	GSE75248	Pure	0.99993526
GSM1947314_3999997114_R02C02	Placenta	-124.66817	-251.60676	Term	Male	GSE75248	Pure	0.99993546
GSM1947396_3999997110_R05C02	Placenta	-145.74215	-225.79535	Term	Female	GSE75248	Pure	0.99994163
GSM1842861_3999984095_R06C02	Placenta	-259.74664	-213.3353	Term	Male	GSE71678	Pure	0.99994347
GSM1843053_3999984068_R05C01	Placenta	-259.01841	-212.97162	Term	Female	GSE71678	Pure	0.9999435
GSM1947403_3999997114_R02C01	Placenta	-124.33151	-283.29597	Term	Female	GSE75248	Pure	0.99994607
GSM1947365_3999997093_R01C01	Placenta	-184.18255	-213.72899	Term	Female	GSE75248	Pure	0.99995131
GSM1842984_3999997059_R06C01	Placenta	-191.69503	-211.84195	Term	Female	GSE71678	Pure	0.99995164
GSM1843127_3999997018_R04C02	Placenta	-188.96801	-213.79907	Term	Female	GSE71678	Pure	0.99995221
GSM1843125_3999997018_R01C02	Placenta	-244.35487	-224.22142	Term	Female	GSE71678	Pure	0.99995584
GSM1947337_3999997114_R01C01	Placenta	-133.53687	-306.27406	Term	Male	GSE75248	Pure	0.99995954
GSM1843043_3999984049_R01C02	Placenta	-380.54335	-339.62987	Term	Female	GSE71678	Pure	0.99995962

GSM1843066_3999984082_R06C01	Placenta	-305.89748	-272.34939	Term	Male	GSE71678	Pure	0.99996254
GSM1843006_3999984005_R06C01	Placenta	-268.64325	-260.51418	Term	Male	GSE71678	Pure	0.9999687
GSM1843088_3999984151_R06C01	Placenta	-274.25497	-267.80432	Term	Male	GSE71678	Pure	0.99997056
201414140060_R04C01	Placenta	-231.86492	-255.09602	First	Male	GSE131945	Pure	0.99997108
GSM1843052_3999984068_R04C01	Placenta	-337.6373	-325.02733	Term	Male	GSE71678	Pure	0.99997533
GSM1947343_3999997114_R03C01	Placenta	-189.91188	-287.09074	Term	Male	GSE75248	Pure	0.99997656
GSM1843030_3999984038_R06C01	Placenta	-211.13832	-292.02348	Term	Female	GSE71678	Pure	0.99998014
GSM1843007_3999984005_R01C02	Placenta	-320.5112	-325.05636	Term	Male	GSE71678	Pure	0.99998026
GSM1843087_3999984151_R05C01	Placenta	-227.26932	-296.31909	Term	Female	GSE71678	Pure	0.99998187
GSM1843038_3999984049_R02C01	Placenta	-254.25909	-298.78789	Term	Male	GSE71678	Pure	0.9999822
GSM1947386_3999997114_R05C01	Placenta	-187.5648	-366.72282	Term	Female	GSE75248	Pure	0.99998755
GSM1843032_3999984038_R02C02	Placenta	-286.08022	-398.25845	Term	Male	GSE71678	Pure	0.99999432
GSM1947351_3999997114_R04C01	Placenta	-283.72784	-504.97028	Term	Female	GSE75248	Pure	0.99999827

References

1. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-9.
2. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics. 2017;33(24):3982-4.
3. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome biology. 2016;17(1):208.
4. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic acids research. 2017;45(4):e22-e.
5. Zhi D, Aslibekyan S, Irvin MR, Claas SA, Borecki IB, Ordovas JM, et al. SNPs located at CpG sites modulate genome-epigenome interaction. Epigenetics. 2013;8(8):802-6.
6. Bibikova M, Le J, Barnes B, Saedinia-Melnyk S, Zhou L, Shen R, et al. Genome-wide DNA methylation profiling using Infinium® assay. Epigenomics. 2009;1(1):177-200.
7. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011;98(4):288-95.

8. Heiss JA, Just AC. Identifying mislabeled and contaminated DNA methylation microarray data: an extended quality control toolset with examples from GEO. Clinical epigenetics. 2018;10(1):73.

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

10. Rosner B, Grove D. Use of the Mann–Whitney U-test for clustered data. Statistics in medicine. 1999;18(11):1387-400.

Statement of Authorship

Title of Paper	Quality control measures for placental sample purity in DNA methylation array analyses
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Qianhui Wan, Shalem Yiner-Lee Leemaqz, Stephen Martin Pederson, Dylan McCullough, Dale Christopher McAninch, Tanja Jankovic-Karasoulos, Melanie Denise Smith, Konstantinos Justinian Bogias, Ning Liu, James Breen, Claire Trelford Roberts and Tina Bianco-Miotto. Quality control measures for placental sample purity in DNA methylation array analyses. Placenta 2019, 88:8-11. PMID: 31569011; DOI: 10.1016/j.placenta.2019.09.006

Principal Author

Name of Principal Author (Candidate)	Qianhui Wan		
Contribution to the Paper	Performed the data analyses and wrote the manuscript.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	23/03/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Shalem Yiner-Lee Leemaqz		
Contribution to the Paper	Assisted with statistics and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Stephen Martin Pederson		
Contribution to the Paper	Assisted with statistics and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Dylan McCullough		
Contribution to the Paper	Sample collection and laboratory work.		
Signature		Date	23/03/20

Name of Co-Author	Dale Christopher McAninch		
Contribution to the Paper	Laboratory work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Tanja Jankovic-Karasoulos		
Contribution to the Paper	Laboratory work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Melanie Denise Smith		
Contribution to the Paper	Assisted in coding and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Konstantinos Justinian Bogias		
Contribution to the Paper	Assisted in coding and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Ning Liu		
Contribution to the Paper	Assisted in coding.		
Signature		Date	23/03/20

Name of Co-Author	James Breen			
Contribution to the Paper	Conceived the study design, supervised the work, assisted in coding and edited the manuscript.			
Signature			Date	23/03/20

Name of Co-Author	Claire Trelford Roberts			
Contribution to the Paper	Conceived the study design, supervised the work and edited the manuscript.			
Signature			Date	23/03/20

Name of Co-Author	Tina Bianco-Miotto			
Contribution to the Paper	Conceived the study design, supervised the work and edited the manuscript.			
Signature			Date	23/03/20

Please cut and paste additional co-author panels here as required.

4 DNA methylation profiling of human placenta samples across early gestation

Abstract

DNA methylation is involved in developmental processes such as X chromosome inactivation and imprinting, as well as tissue-specific gene expression. However, the function of DNA methylation in the human placenta during early development, including trophoblast proliferation and differentiation, is not well understood. We do know the placenta is hypomethylated compared to other human tissues and contains partially methylated domains [1]. We profiled the methylome of 125 placental chorionic villous samples (excluding outliers) across 6 to 23 weeks' gestation using Illumina HumanMethylationEPIC arrays. At approximately 10 weeks' gestation, maternal blood starts to flow into the placenta after which placental oxygen tension rises. Samples up to and after 10 weeks' gestation were compared and 295 differentially methylated regions (DMRs) were identified between these two groups. Partially methylated domains (PMDs) and placenta specific imprinting control regions (ICRs) were established from 6 weeks' gestation and were stable across early gestation. Placental promoters and enhancers were overall hypomethylated across early gestation compared to non-placental tissues. In addition, we identified 91 DNA methylation sites that could be used for predicting gestational age. This study represents a characterisation of placental DNA methylation profile which has important implications for the aetiology of pregnancy complications related to placental dysfunction.

Keywords: Human placenta, DNA methylation, early gestation

4.1 Introduction

It is well known that the human placenta plays crucial roles in fetal development and in maintaining a healthy pregnancy for both mother and child. The development of the fetus depends on the exchange of gases, nutrients and wastes with the maternal circulation through the vasculo-syncytial membrane of the placenta [2]. Moreover, it is well-known that peptide and steroid hormones and cytokines released by the placenta have vital effects on both maternal and fetal physiology. Importantly, placenta secreted factors orchestrate maternal physiological adaptations to pregnancy without which pregnancy cannot proceed. The placenta dynamically differentiates across gestation to enable its expanding functional capacity while also responding to maternal environmental exposures that modulate its function [2].

DNA methylation is an epigenetic modification that characterised by DNA methyltransferase addition of a methyl group predominantly at the 5th position of cytosines in mammals [3]. DNA methylation has important functions during fetal and placental development. Methylation of cytosines predominantly occurs at CpG dinucleotides but they are also found at non-CpG sites (i.e. CHH, including CpA, CpT, and CpC). Interestingly, only some specific cell types such as pluripotent stem cells and oocytes have non-CpG methylation [4]. Canonically, DNA methylation at regulatory regions such as promoters is associated with repression of gene expression [5]. DNA methylation plays an important role in imprinting of genes and the inactivation of one X chromosome in females [6].

Although DNA methylation in term placenta and that from complicated pregnancies have already been studied [7], the pattern of DNA methylation across early gestation is not as well characterised. It is known that the human term placenta has a unique DNA methylation profile and is hypomethylated compared to other healthy tissues [8]. It has also been shown that DNA methylation levels increase from first, to second and again to third trimesters [9]. Comparison of first trimester (8-10 weeks' gestation) with late first trimester villous cytotrophoblasts (12-14 weeks' gestation) identified large regions of partially methylated DNA called partially methylated domains (PMDs) [1] and more methylation in late vs early first trimester cytotrophoblasts [10]. This coincides with the transition from a state of relative hypoxia in the placenta to an oxygen tension that equates with that in arterioles following commencement of maternal blood flow into the intervillous space that starts to occur at approximately 10 weeks' gestation [11]. Currently, there are no studies that have used EPIC arrays to comprehensively characterise human placenta tissue across 6-23 weeks' gestation, so the aim of this study was to profile placental DNA methylation changes across the first half of gestation.

4.2 Methods

4.2.1 Ethics statement

Prior to collection of placental tissue samples, written, informed consent was obtained from all subjects involved in this study. Collection of placental tissue from elective termination of first and second trimester pregnancies was approved by the Queen

Elizabeth Hospital and Lyell McEwin Hospital Human Research Ethics Committee (TQEH/LMH/MH and HREC/16/TQEH/33).

4.2.2 Sample Description

Sixty-three first trimester and sixty-eight second trimester placentas were obtained from elective pregnancy terminations from the Pregnancy Advisory Centre at Woodville, South Australia. Gestational age was determined using transabdominal ultrasonography. Women with infection, endocrine abnormalities, antiphospholipid syndrome or other known complications were excluded from the study. Placentas were collected and dissected within 15 minutes of termination and placental villous tissues were isolated and immersed in RNA Later solution (Invitrogen) and frozen at -80°C for subsequent DNA and RNA extraction. The meta data for all samples is listed in Supplementary Table S4.1.

4.2.3 Array processing

DNA was extracted from 131 placental chorionic villous tissue samples using a modified version of the TES protocol (10mM Tris(T), 2mM Na_2EDTA(E), 10% SDS(S)) [12]. For each sample, 1µg of DNA was sent to PathWest Laboratory Medicine (QEII Medical Centre, Perth, Western Australia) for bisulfite-conversion and hybridisation to the Illumina Infinium Methylation EPIC BeadChips according to the manufacturer's instructions.

4.2.4 RNA sequencing

RNA was extracted from samples using Trizol according to the manufacturer's instruction. Illumina® TruSeq® Stranded Total RNA Sample Preparation kits were used for preparing sequencing libraries. Adapters were added to the end of each sequence and the sequencing by synthesis process will be accomplished using Illumina HiSeq 2000 system at the Australian Cancer Genomics Facility in Adelaide. The result of the sequencing will be aligned to the human genome and the counts of the transcripts will be calculated for later bioinformatic analysis.

4.2.5 EPIC array data analysis

4.2.5.1 Quality control

The quality of the EPIC array data was assessed using *minfi* package [13]. Scatterplots of median unmethylation (Unmeth) signal versus median methylation (Meth) signal values were generated for assessing quality for each sample. Good samples clustered together, while failed samples tended to separate and had lower median intensities (log median intensity < 11 were considered as poor quality samples) [13]. In addition, the sample-dependent control probes in EPIC array were used to check the quality of bisulfite conversion for type I and II assays, non-polymorphic performance, specificity of matching type I and II probes and the system background.

Principle component analysis (PCA) was used for checking sample outliers. PCA was used to summarise matrix data into fewer dimensions by fitting the matrix to orthogonal

axes [14]. Except for outliers, we also checked the intensity of probes on X and Y chromosomes using *minfi* and *ewastools* package, to check fetal sex for placental samples. Six samples that were not pure placenta samples were excluded after a quality control step documented in Chapter 3 [15].

4.2.5.2 Probes on microarray and filtering of failed probes

Illumina Infinium Human MethylationEPIC array contains 866238 probes. When related to CpG islands, locations of these probes are categorised as in CpG island, N-shelf, N-shore, S-shelf, S-shore and open sea. When related to genes, these probes are located in gene body, intergenic region, 5'UTR, 3'UTR and regions around transcription start sites (TSS) (including TSS1500, TSS200 and 1st Exon). There are also probes detecting non-CpG methylation sites (CHH sites). These CHH sites are based on CHH methylation identified in human stem cells. Comparing with Illumina 450K array, Illumina EPIC array has increased probe numbers especially in open sea regions (probe positions relative to CpG islands) and gene body (probe positions relative to genes).

Failed and unwanted probes were filtered. The failed probes removed including 19431 probes with detection $P > 0.01$, 13490 probes with probes < 3 beads in 5% of the 96 samples, 25836 probes with SNPs at CpG/SBE sites on probes [13], and 42249 cross-reactive probes [16]. In favour of downstream analysis, the 15321 probes on X/Y chromosomes and 133603 probes with SNPs at probe body (not located at CpG/SBE sites) were also removed to decrease the disturbance variables. In total, 249472 probes were removed, and 616766 probes remained for following processing.

4.2.5.3 Background and dye bias correction

Filtered data were pre-processed with *ENmix* package to remove background noise and correct dye bias [17]. *ENmix* method modelled methylation signal intensities with a exponential-normal mixture distribution, and the background noise is modelled with a truncated normal distribution. The background normal distribution parameters were estimated using out-of-band (OOB) intensities which were proved to be better than internal negative controls [17, 18]. Regression on logarithm of internal control probes (RELIC) from *Enmix* package was used for dye bias correction to improve accuracy of methylation beta value estimates [19], which was also implemented in *ewastools* package as a dye bias correction tool.

4.2.5.4 Data normalisation

Similar to the Illumina 450K platform, Infinium I (type I probes) and Infinium II (type II probes) assays were used in the Illumina EPIC platform to assess the status of DNA methylation for more than 850,000 cytosines distributed over the human genome. Normalisation is needed since the data generated by the two chemical assays are not entirely compatible, which could cause potential bias in the DNA methylation analyses [16, 20]. The corrected data were normalised with Beta-mixture quantile (BMIQ) normalisation method implemented in *ChAMP* package, which adjusted beta-values of type II probes based on the statistical distribution of type I probes to correct probe design bias [21]. The normalised data was used for downstream analyses.

4.2.5.5 Identification of differentially methylated probes (DMPs)

Beta values (β) and M values are used to analyse differential methylation. Beta values ($\beta = \frac{Meth}{Meth+Unmeth+100}$) are the proportion of methylation intensity (Meth) over the total intensity of methylation and unmethylation (Unmeth). M values are logit transformed beta values ($\mathrm{M} = \log 2\left(\frac{\beta}{1-\beta}\right)$) and were used for analysing differentially methylated probes across early gestation [22, 23]. Batch effects were adjusted in the linear models and relative quality weights for samples were estimate based on sample variances in groups with highly variable samples [24] downweighted in the models [25]. Differentially methylated probes were identified with false discovery rate (FDR) less than 0.05 and the change of Beta values ($|\Delta\beta| > 0.2$) higher than 0.2 between groups.

4.2.5.6 Identification of differentially methylated regions (DMRs)

DMRcate package in R was used to identify of DMRs [26]. Briefly, results from linear modelling adjusted for technical variance were analysed using a Gaussian kernel smoother (bandwidth of 1000 bp and scaling factor of 2) to regions with differential DNA methylation. Through clustering significant DNA methylation sites located within 1000 bp windows that contains three or more DNA methylation sites, DMRs were assigned. Significance of each DMR was measured by Stouffer's p-value that is the combined p-value for all sites in a region calculated using Stouffer's method. Changes of mean Beta values higher than 0.2 ($|\text{mean}\Delta\beta| > 0.2$) between groups were considered as biologically meaningful DNA methylation changes [27, 28]. The gene

ontology (GO), KEGG pathway analysis and gene set analysis (GSA) were conducted using the *missMethyl* R package [29].

4.2.5.7 Analyses of RNA sequencing data

Quality control of the raw RNA sequencing data were applied with *FastQC* [30]. Samples that passed the quality control were used for trimming sequence adapters by *AdapterRemoval* [31]. Then trimmed sequences were aligned to human genome hg19 to keep consistent with EPIC array data using RNA mapping program *STAR* [32] and duplicates removed by *picard* [33]. Counts of each gene were quantified using *featureCounts* [34]. Differential expression analysis was performed using *limma* with normalised counts generated using TMM (trimmed mean of M-values) [35] method in *EdgeR* [36] package. Gene expression levels are presented by log transformed counts per million (lcpm).

4.2.5.8 Gestational age analysis

Normalised beta values of 125 samples (6 outliers excluded) were transformed to M values for gestational age (GA) analysis. We filtered out 133588 non-variant probes that are detected by *caret* R package with beta values of 95% of the samples with zero variance. Six placenta samples were excluded from this analysis due to apparent contamination with other tissues [15]. Importance of correlations between probes and GA were determined for 482720 probes using recursive feature elimination (RFE) method which is a feature selection method from *caret* package [37]. The top 50% of the 482720 probes that most correlated with GA according to the RFE methods were

selected for downstream analysis for predicting GA. Using the R package *glmnet*, elastic net regression was performed to select probes predictive of GA. According to Horvath's study [24], the parameter alpha of the elastic net model was set to 0.5 allowing the equal contribution of the ridge and lasso methods. The lambda parameter of the model was chosen according to 10-fold cross-validation of the training data (80% of 125 samples). The predicted values of GA were generated using the training coefficient values and probes selected from this regression and the accuracy of predicted GA was assessed by correlation coefficients between DNA methylation age and chronological age.

4.2.5.9 Analysis of partially methylated domains

In order to identify partially methylated domains, we filtered out probes on promoter regions and CpG islands that were known as low methylated regions according to a previous study [1]. Partially methylated domains from Schroeder et al. were used to verify and identify changes within these domains. To investigate the DNA methylation level of regions such as promoters and enhancers within and out of PMDs, regions were separated into two groups (i.e. in PMDs and not in PMDs) and the DNA methylation between groups is compared. Annotations of promoters and enhancers were acquired from Fantom5 database (http://slidebase.binf.ku.dk/).

4.2.5.10 Analysis of DNA methylation at imprinting control regions

In order to identify the change of DNA methylation at imprinted genes, probes in imprinting control regions (ICRs) were used according to the annotation of EPIC array

(http://zwdzwd.github.io/InfiniumAnnotation#current). ICRs with mean beta values less than 0.25 or more than 0.75 across all the 125 placenta samples were filtered out. Placenta specific ICRs (13 in total) used in this paper were identified by Court et al. [38]. Using linear models in *limma* package, differentially methylated ICRs were identified with false discovery rate (FDR) less than 0.01 and the change of Beta values higher than 0.2 ($|\Delta\beta| > 0.2$) between groups.

4.2.5.11 Annotation

EPIC microarray manifest and annotation of all CpG sites within these analyses was performed using Bioconductor packages *IlluminaHumanMethylationEPICmanifest*, *IlluminaHumanMethylationEPICanno.ilm10b4.hg19*. Bioconductor/R package *BSgenome.Hsapiens.UCSC.hg19* and *biomaRt* were used for annotation of DMR overlapped genes. Annotations of promoters and enhancers were acquired from Fantom5 database (http://slidebase.binf.ku.dk/).

4.3 Results

4.3.1 The placenta is hypomethylated and DNA methylation increased from early to mid-gestation

In order to gain more insight into the placental DNA methylation profile in early to mid-gestation, we used 125 human placenta samples from 6-23 weeks' gestation (Supplementary Figure S4.1). We excluded 6 placenta samples that were previously

shown to be placental villous tissue mixed with decidua [39]. PCA using transformed percentage of DNA methylation (M values) showed that the placenta samples clustered based on gestational age (Figure 4.1A). Of 616477 probes after filtering, there were more hypomethylated probes than hypermethylated probes with methylation percentage (Beta values) ranging from 0 to 1, respectively (Figure 4.1B). DNA methylation increased across gestation (GSE74738, Supplementary Figure S4.2). The top 1% of significantly changing probes contributing to PC1 showed gradual increases of DNA methylation with increasing gestation (Figure 4.1C). Promoters and enhancers, especially placental promoters and enhancers as identified by the FANTOM5 project, had overall low levels of methylation (Figure 4.1D).

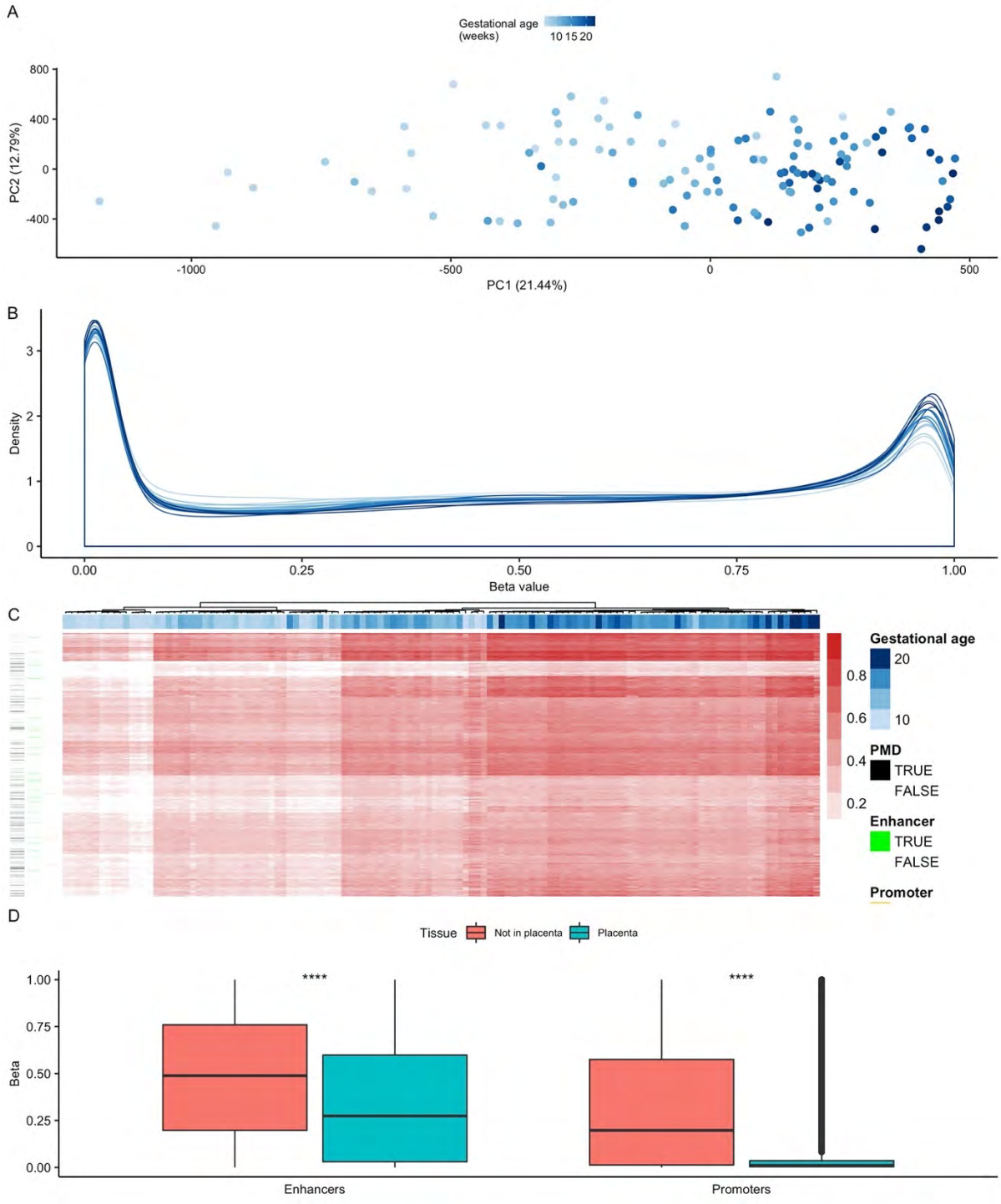

Figure 4.1 DNA methylation patterns in human placental chorionic villous samples across 6-23 weeks' gestation. (A) PCA using M values shows a cline across early gestation. (B) Density plots for all 125 samples based on all probes shows greater hypomethylation than hypermethylation peaks. (C) Heatmap of the top 1% of differentially methylated probes contributing to PC1, DNA methylation (Beta values) increased across early gestation. (D) DNA methylation at promoters and enhancers annotated in non-placental tissues (orange, FANTOM5 data) compared to that at placental (green) promoters and enhancers.

4.3.2 DNA methylation increased at the promoter region of solute carrier family 2 member 3 (*SLC2A3*)

SLC2A3 was chosen as an example of negatively correlated DNA methylation and gene expression. We used both RNA sequencing data and DNA methylation data to investigate correlations between DNA methylation and gene expression (Supplementary Figure S4.3). Our results showed that DNA methylation of *SLC2A3* promoter increased across gestation (Figure 4.2A) and gene expression decreased across gestation (Figure 4.2B). The DMR that overlapped with the promoter of SLC2A3 gene also overlapped with transcription factor binding motifs including motifs of KLF3, SP2, ETS1, RUNX1 and PRDM15 that annotated according to FATOM 5 data [40, 41] (Figure 4.2C).

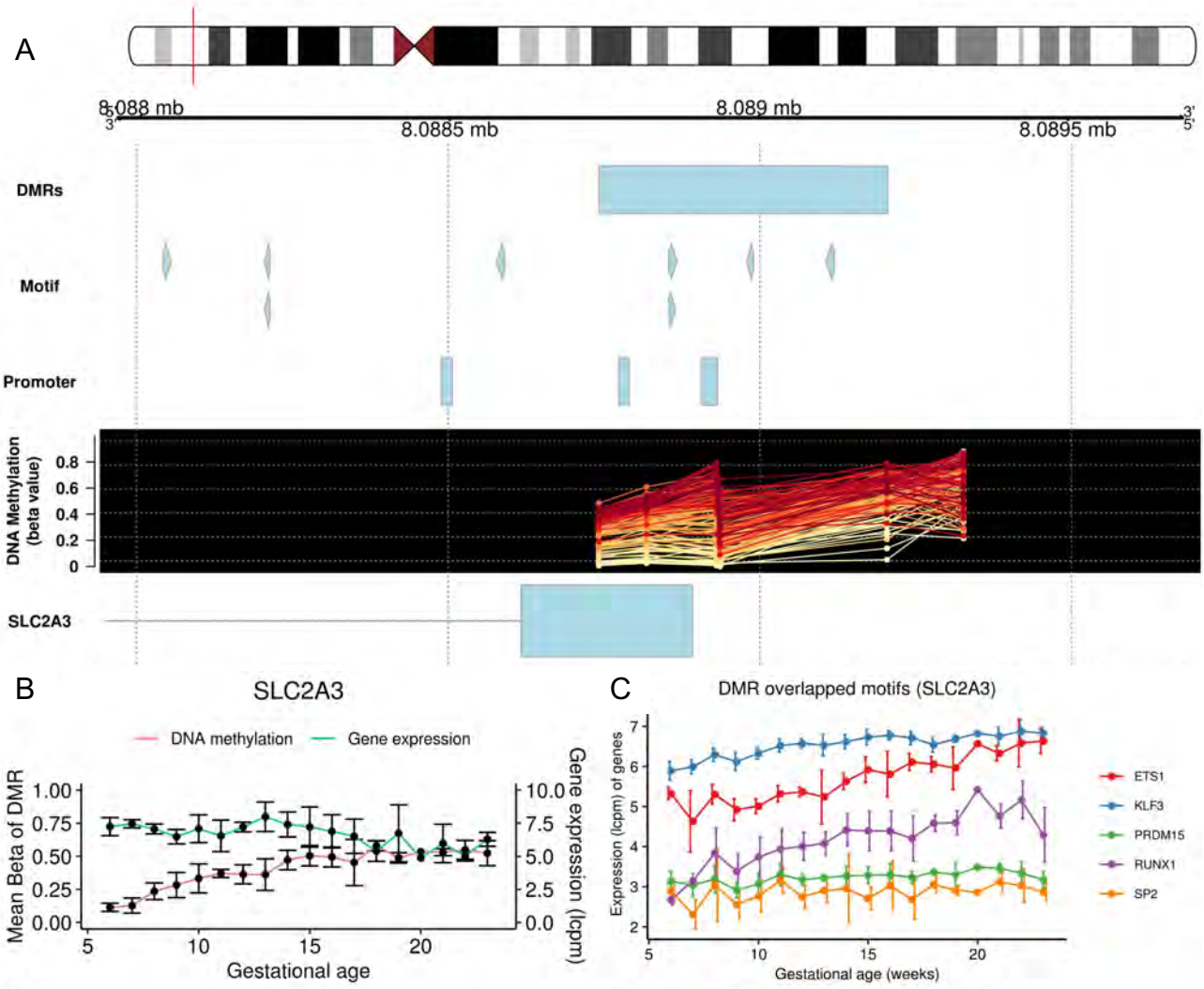

Figure 4.2 DNA methylation of *SLC2A3* gene promoter on Chromosome 12. (A) DNA methylation of the DMR overlapped with *SLC2A3* gene promoters (yellow->red: 6-23 weeks' gestation). The DMR contains 5 DNA methylation sites detected by EPIC probes. (B) Mean beta values (% DNA methylation) of DMR increased and then plateaued across early gestation while the RNA expression of *SLC2A3* decreased and plateaued across early gestation. The error bar represents the standard deviation for each gestational week. (C) The expression of genes that encode transcription factors (ETS1, KLF3, PRDM15, RUNX1 and SP2) with their motifs identified within the DMR.

4.3.3 Partially methylated domains (PMDs) are stable from 6-23 weeks' gestation

Previous studies have shown that the placenta DNA methylome is characterised by PMDs [1]. We assessed DNA methylation at published human placenta PMDs across

our 125 first and second trimester samples. We have confirmed that PMDs exist in the human placenta and show that PMDs are stable from 6-23 weeks' gestation (Figure 4.3A). Probes in CpG islands and promoters were removed before analysing DNA methylation in PMDs because the DNA was unmethylated in those regions [42]. As expected, the expression of genes in PMDs was repressed compared to genes located in non-PMDs (Figure 4.3B). To investigate the distribution of DNA methylation within PMDs in placenta, DNA methylation level at promoters, enhancers and gene body regions were analysed. DNA methylation of promoters in PMDs were more variable than promoters that were not in PMDs. DNA methylation at placenta specific enhancers did not differ between those located in PMDs or non-PMDs. However, DNA methylation in the gene bodies was decreased dramatically in PMDs compared to non-PMDs (Figure 4.3C). Numbers of probes in PMDs and non-PMDs are listed in Supplementary Table S4.2. These placenta PMDs were established early from 6 weeks' gestation and remain consistently partially methylated across early to mid-gestation (Supplementary Figure S4.4). The GO for genes in PMDs showed that these genes were related to sensory perception of smell and humoral immune response (Supplementary Table S4.3). Though the DNA methylation is overall stably partially methylated in PMDs, there are 10 DMRs (when comparing placenta up to and after 10 weeks' gestation) in PMDs. GO analysis of the genes overlapped with these DMRs showed that the related genes were enriched in mitochondria genome maintenance, ion channel and reproduction. There are six DMRs in the promoter regions of genes including *LOC105376599*, *CALCR*, *KCNK2*, *KCNQ5-IT1*, *LINC01845* and *RPTN.* Interestingly, DMR related genes including *LOC105376599* (*ANO3-AS1*), *KCNQ5-IT1*, *ZBTB20-AS1*, *MIR548* and *LINC01845* were non-coding genes that could be

associated with ion and glucose homeostasis in placenta tissue across early to mid-gestation (Supplementary Table S4.4).

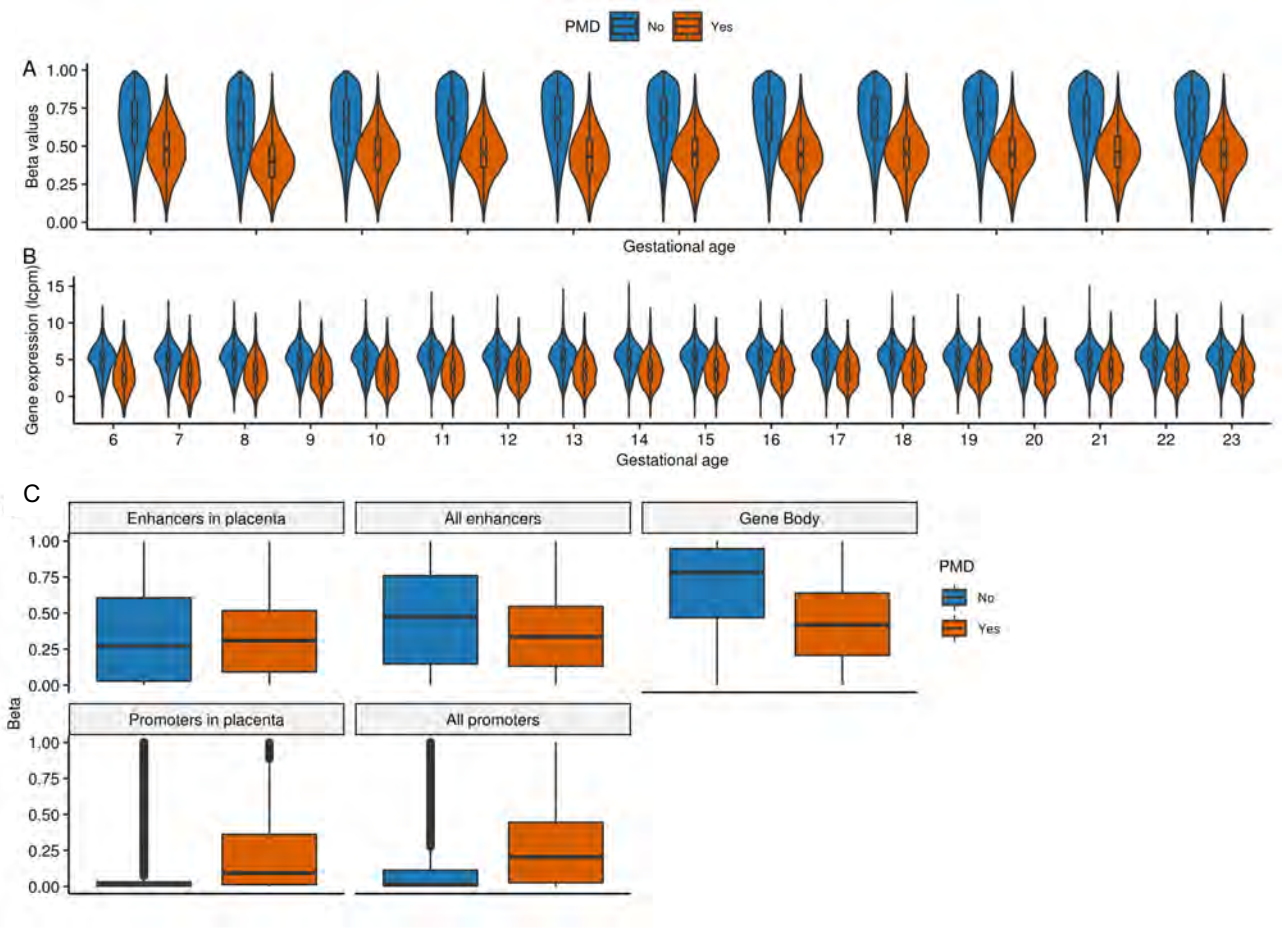

Figure 4.3 DNA methylation and gene expression of PMDs are stable from 6-23 weeks' gestation. (A) DNA methylation and (B) gene expression of genes in PMDs. Expression of genes in PMDs was consistently repressed compared to genes that are not in PMDs. (C) DNA methylation of gene body, enhancer and promoter regions in PMDs.

4.3.4 Gene imprinting across early gestation

In order to study placental DNA methylation of imprinted genes from 6-23 weeks' gestation, the imprinting control regions (ICRs) were investigated. Placenta-specific ICRs (11 in total) were consistently 50% methylated across early to mid-gestation (Figure 4.4A). Within 727 ICRs (mainly overlapping CpG islands), more than 50% of

these ICRs were consistently partially methylated across early gestation (Figure 4.4B). 61 ICRs associated with 22 genes were located in PMDs. Details of these ICRs are listed in Supplementary Table S4.5.

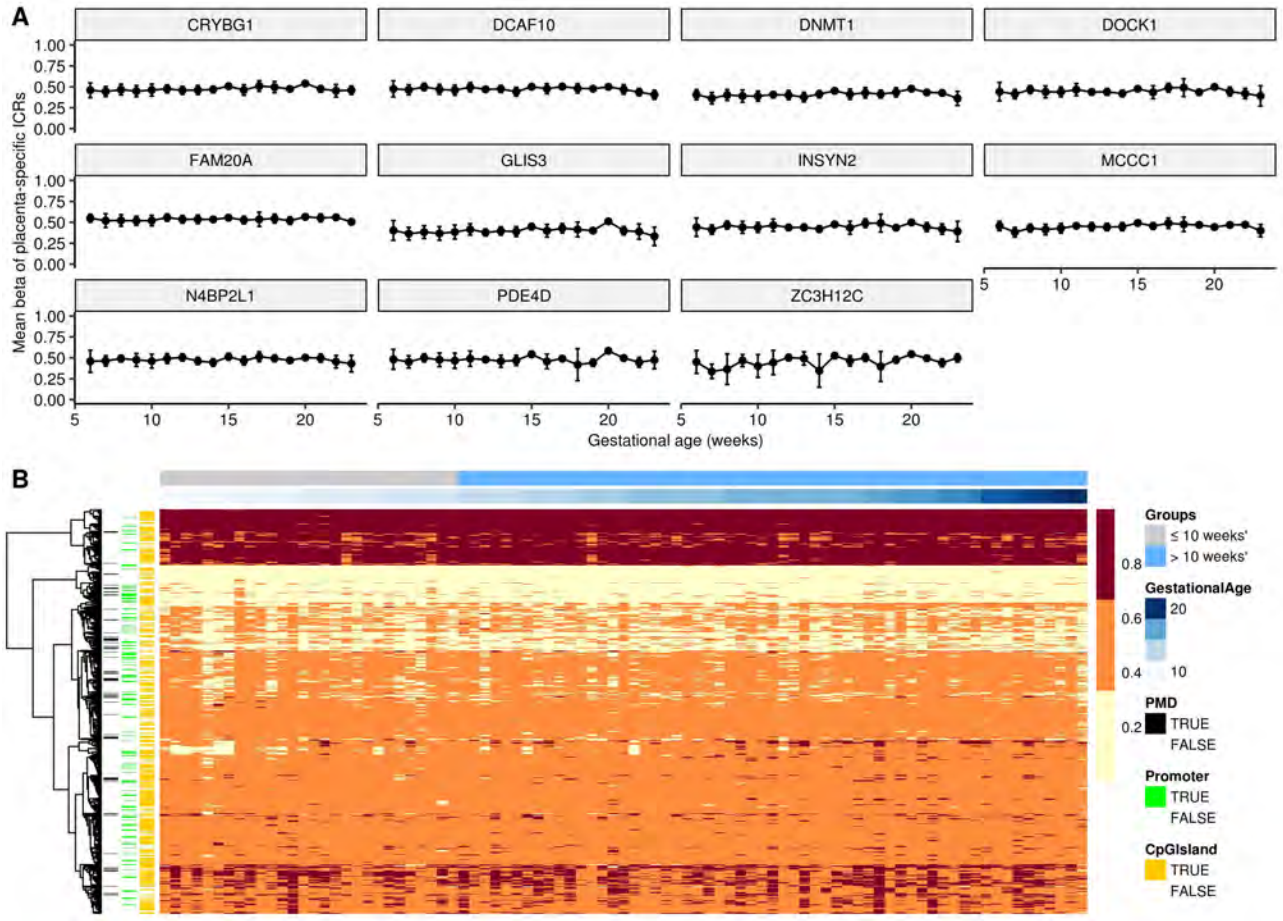

Figure 4.4 DNA methylation of imprinted genes. (A) The placenta-specific imprinting control regions were partially methylated across early to mid-gestation. (B) DNA methylation of all ICRs annotated in the EPIC array.

4.3.5 DNA methylation changes in the placenta up to and after 10 weeks' gestation

Maternal blood flow into the intervillous space commences at approximately 10 weeks of gestation, so we assessed differential DNA methylation between groups ≤ 10 weeks'

(n= 42) and > 10 weeks' (n=83) gestation. Cohort clinical information of these sample groups is listed in Supplementary Table S4.6. Probes in ±2000bp from transcription start sites (TSS2000), promoter and enhancer regions were used for identifying differentially methylated regions (DMRs), respectively. Gene ontology of 295 DMRs (identified using all sites) showed that the genes associated with these DMRs are involved in biological processes such as angiogenesis and cell migration (Supplementary Figure S4.5). Using probes in TSS2000, and after adding a threshold of P-values <0.05 and beta > 0.2, we identified 42 DMRs that showed significant increases of DNA methylation in the >10 weeks' gestation group and 1 DMR showed a significant decrease of DNA methylation in this group (Figure 4.5 A-B). RNA sequencing data showed that there was not always a correlation between DNA methylation and gene expression (Figure 4.5 C). Increased DNA methylation in DMRs that overlapped with TSS2000 of *CCR7*, *IL17D* and *EPB42* were correlated with decreased gene expression. Details of all DMRs are listed in Supplementary Table S4.7.

Besides analysing differentially methylated regions between groups ≤ 10 weeks' (n= 48) and > 10 weeks' (n=83) gestation, we also compared the large differentially methylated blocks (DMBs) between these two groups. DMBs are large genomic regions varying from 10kb to megabases that contain hundreds of DNA methylation sites [43]. We identified 25 statistically significant DMBs and these were often located in non-PMDs or in regions of the genome where there are sparse PMDs. This could be due to the fact that PMDs are stably partially methylated across early to mid-gestation (Supplementary Figure S4.6).

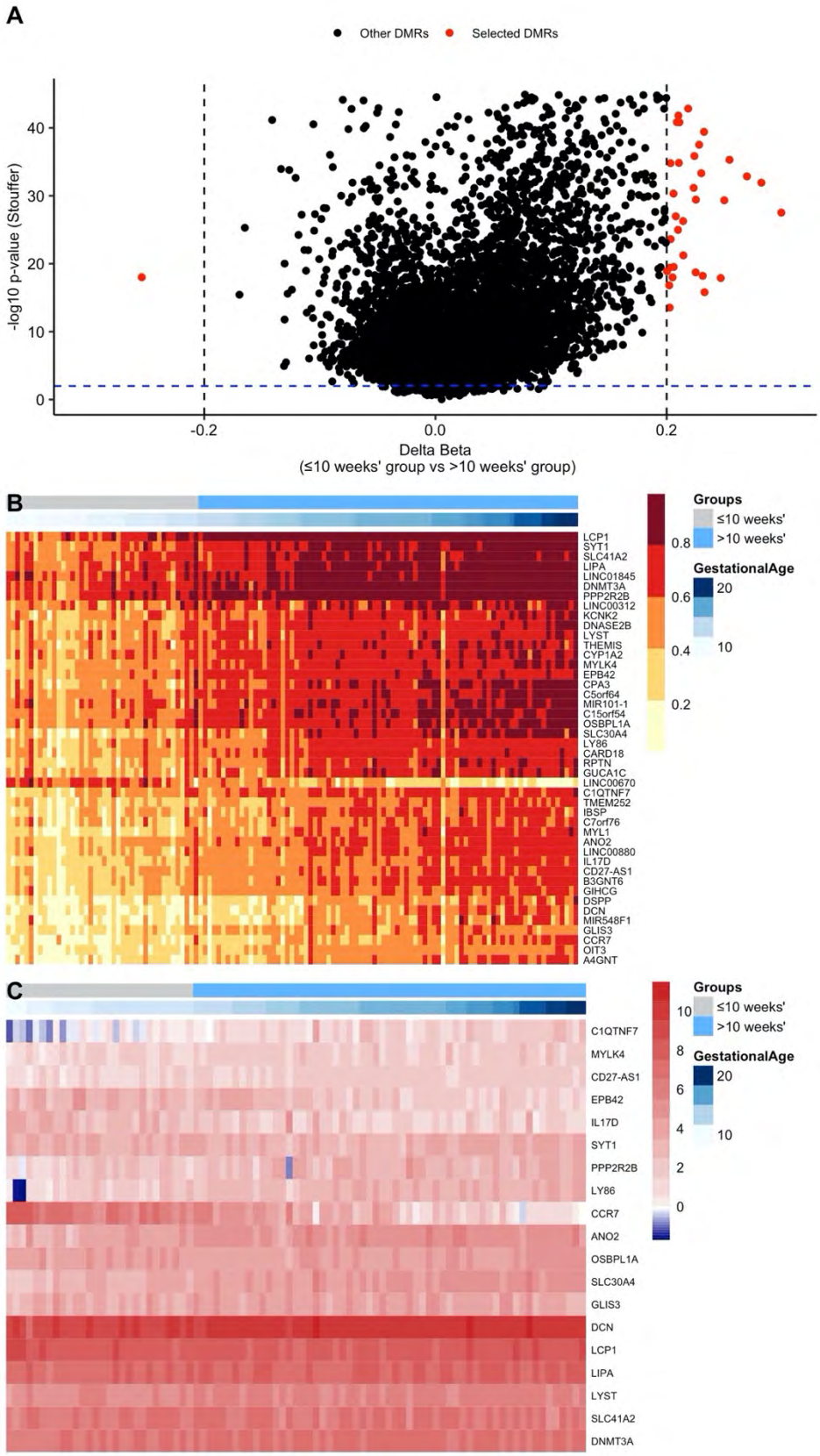

Figure 4.5 Placental DNA methylation changed across early gestation when comparing ≤ 10 vs > 10 weeks' gestation. (A) Volcano plot of all regions identified using probes in TSS2000. Thresholds of differentially methylated regions (DMRs) were applied with p-value < 0.05 and absolute ΔBeta >0.2. (B) Heatmap shows the

change of DNA methylation of DMRs identified using probes in TSS2000. Gene names labelled represent the nearest gene located next to the DMR. (C) RNA sequencing data shows gene expression of genes overlapped with DMRs identified using probes in TSS2000, Genes with low expression were filtered out.

4.3.6 DNA methylation of C-C motif chemokine receptor 7 (*CCR7*)

CC-chemokine receptor 7 (CCR7) is related to both immunity and immune tolerance. We studied this gene because *CCR7* has only 1 main isoform which was differentially expressed in placenta and the DMR located at the promoter of this gene. Our study showed that DNA methylation of the promoter of *CCR7* increases after 10 weeks' gestation and expression of *CCR7* decreased (Figure 4.6A-B). ETS1 and RUNX1 are 2 transcription factors that bind to the DMR overlapped sequence (Figure 4.6C). The increase of DNA methylation at this DMR may affect binding of these transcription factors.

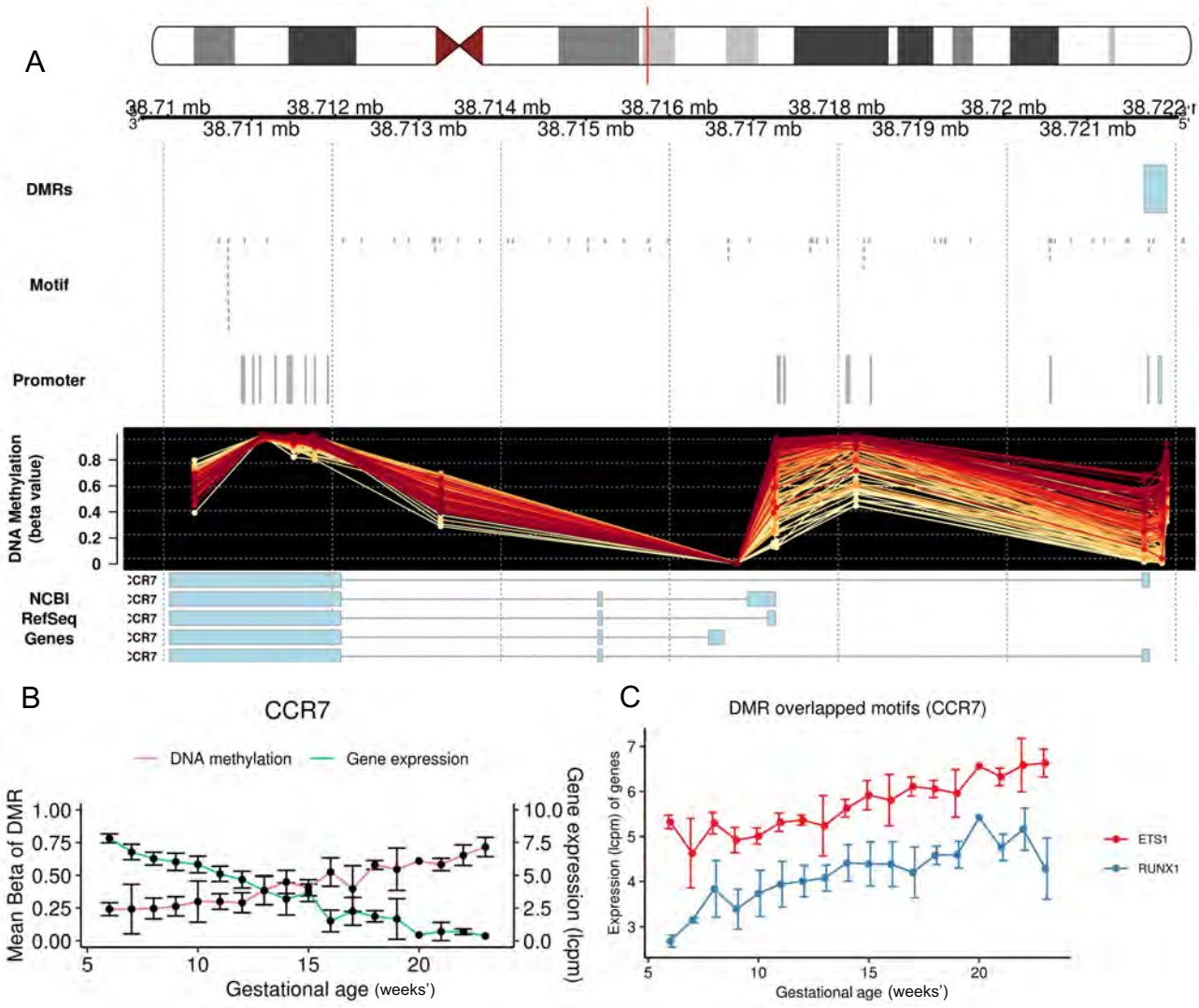

Figure 4.6 DNA methylation of the DMR overlapping *CCR7* gene promoter on Chromosome 17. (A) DNA methylation of *CCR7* gene promoter. This DMR contains 3 DNA methylation sites detected by EPIC probes. Mean beta values (proportion of DNA methylation) of DMR increased across early gestation (yellow->red: 6-23 weeks' gestation) while *CCR7* gene expression decreased over this time (B). (C) The expression of genes that encode transcription factors (ETS1 and RUNX1) with their motifs overlapped with the DMR.

4.3.7 DNA methyltransferase 3 alpha (*DNMT3A*)

As a DNA methyltransferase involved in *de novo* methylation, *DNMT3A* was differentially methylated, which could be associated with DNA methylation involved in developmental processes. Differential methylation analyses between groups ≤ 10

weeks' (n= 48) and > 10 weeks' (n=83) gestation showed that the DMR is in the gene body of *DNMT3A* gene but overlapped with the regulatory region of *MIR1301* (Figure 4.7A). DNA methylation of this DMR increased across gestation but was not correlated with gene expression because gene expression of *DNMT3A* is not changing over early gestation (Figure 4.7B). The expression of MIR13101 could be affected by DNA methylation changes at this DMR which is not included in this study and deserve further investigation. TAL1, ZNF143, STAT6 and ZNF165 are 4 transcription factors that binding to the DMR overlapped sequence (Figure 4.7C). The increase of DNA methylation at DMR may affect the binding of these transcription factors.

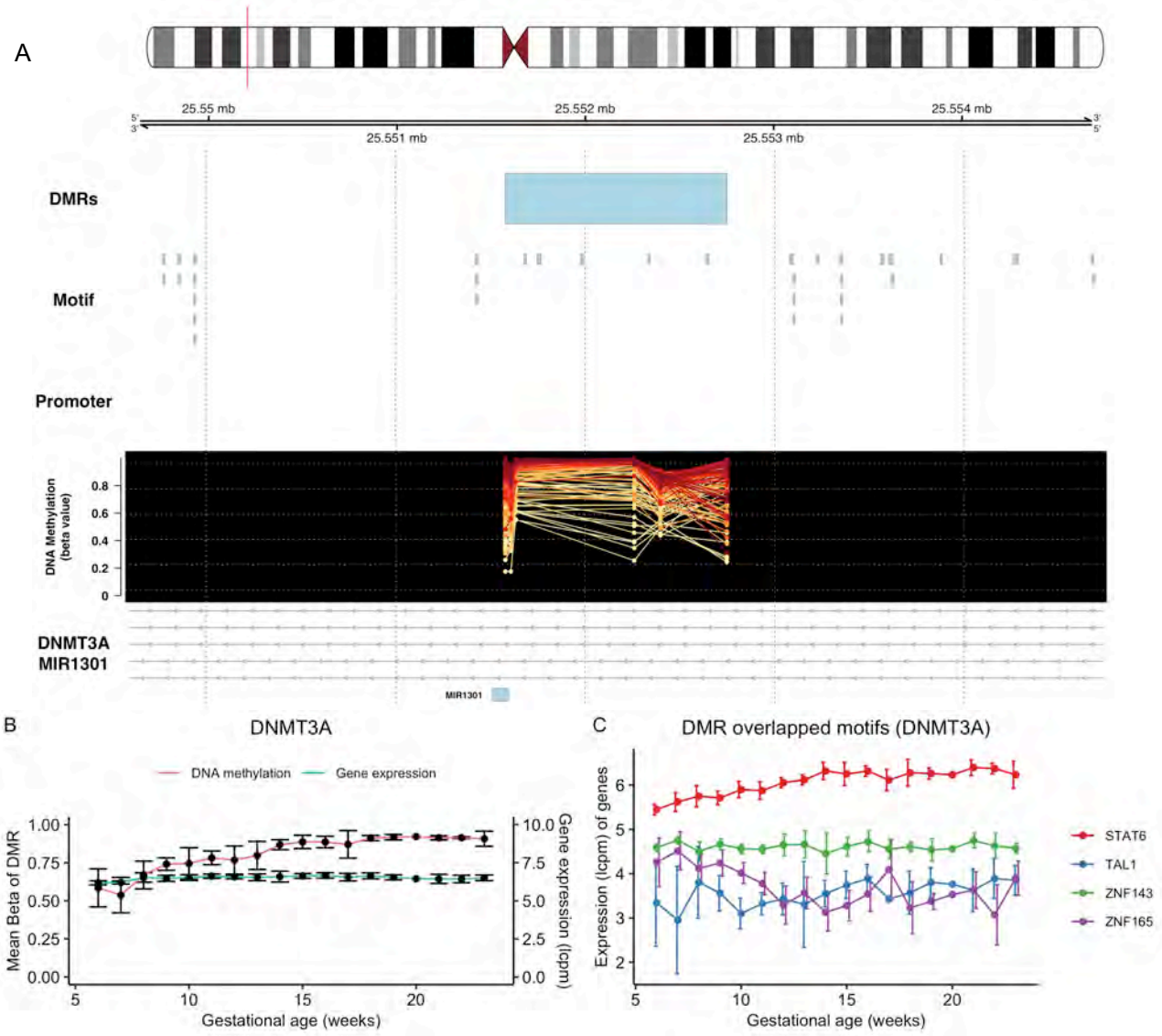

Figure 4.7 DNA methylation of *DNMT3A* on Chromosome 2. (A) DNA methylation of the DMR in the first intron of *DNMT3A*. Mean beta values (proportion of DNA methylation) of DMR increased across early gestation (yellow->red: 6-23 weeks' gestation) while *DNMT3A* gene expression didn't change over this time (B). (C) The expression of genes that encode transcription factors (TAL1, ZNF143, STAT6 and ZNF165) with their motifs overlapped with the DMR.

4.3.8 DNA methylation sites predict gestational age

Since the age of many human tissues and cell types can be accurately predicted using their DNA methylation profiles [24] we used EPIC array data of 125 samples to establish a gestational age (GA) clock. GA of all 125 samples were accurately

predicted (Figure 4.8A-B). There were 91 probes selected for calculating GA. Methylation levels at 41 probes were negatively correlated with GA, while methylation levels at 50 probes were positively correlated with GA (Figure 4.8C). Details of these 91 clock probes are listed in Supplementary Table S4.8.

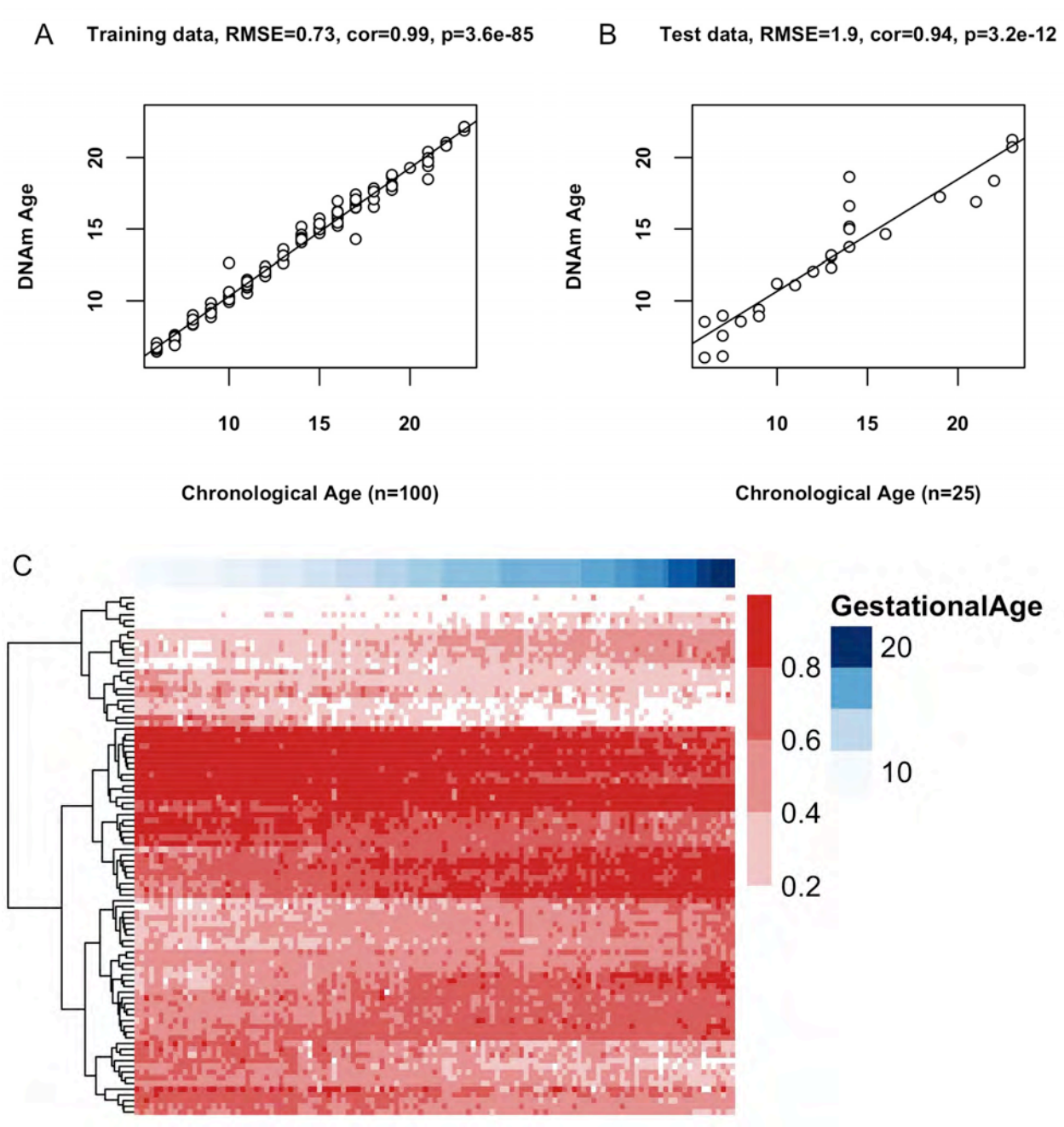

Figure 4.8 Epigenetic clock based on 91 probes. (A) DNA methylation age of training data (n=100, 80% of all samples) and test data (n=25, 20% of all samples). (B) DNA methylation of 41 probes was negatively correlated with GA and that of 50 probes was positively correlated with GA.

4.4 Discussion

In this study, we profiled placenta DNA methylation using 125 placental chorionic villous samples from 6-23 weeks' gestation. We have shown, with much greater temporal detail than previous studies, that the placenta in early gestation is hypomethylated compared to other normal human tissues (FANTOM5 data) [1]. It is known that placental DNA methylation increases across gestation [9] and our results verified this increase in DNA methylation from early to mid-gestation. DNA methylation was more variable in samples of earlier gestational ages compared with later gestational ages. This indicates that the DNA methylation changes across 6-23 weeks' gestation may be associated with the control of placental developmental at the epigenetic level. DNA methylation of DMRs at *SLC2A3*, *CCR7* and *DNMT3A* significantly changes across early to mid-gestation and RNA sequencing confirmed the negative correlation of gene expression of *SLC2A3* and *CCR7* at this time. This study adds to current published studies looking at DNA methylation changes during placental development [10, 44, 45].

Promoter and enhancer regions of placental chorionic villous tissue DNA tend to be consistently hypomethylated across early to mid-gestation. Many promoters are normally hypomethylated in tissues as they usually contain unmethylated CpG islands permitting transcription [46]. It is interesting to note the hypomethylation of enhancers in placental DNA. Although not completely understood, the functions of these enhancers could be associated with the gene regulation in placenta. A study on mouse and rat trophoblasts showed that unlike promoters, enhancers are more species-

specific and less conserved [47]. Whether these enhancers play an important role during the regulation of early placental development will be an interesting area to investigate.

Partially methylated domains (PMDs) are characteristic of a subset of hypomethylated cells such as colon and breast cancer cells and are also found covering 37% of the placental genome [1]. The PMDs of placental DNA, along with the invasive behaviour of extravillous trophoblasts both resemble attributes of cancer [48]. Schroeder et al. has also shown that similar to cancer cells, PMDs are constantly partially methylated in first trimester and term placenta samples [1, 10, 49]. Using 367159 probes after filtering out probes on CpG islands, CpG island shores and promoters, we confirm the existence of PMDs and show they are stable in the placental genome from 6 to 23 weeks' gestation. Also, DNA methylation and gene expression at these PMDs were low compared with other regions. Our data showed that ICRs were stably 50% methylated, especially for placenta specific ICRs, across early to mid-gestation. Though the DNA methylation was approximately 50% in PMDs, more than 95% of ICRs were not in PMDs indicating the function of PMDs is likely not related to genomic imprinting. The biological functions of these PMDs still need to be further investigated.

The most novel analyses in this study revealed DNA methylation changes between samples collected up to and after initiation of maternal blood flow to the placenta (≤ 10 vs > 10 weeks' gestation). The observed DMRs in promoter regions could be related to the immune response during early pregnancy as the DMRs overlapped with immune related genes (*CCR7*, *CD27*, *THEMIS*) [50]. Genes related to ion channels and transporters such as *ANO2*, *SLC30A4* and *SLC41A2*, placenta-specific imprinted

gene *GLIS3* [51] and DNA methyltransferase *DNMT3A,* were differentially methylated with increased DNA methylation in the > 10 weeks' gestation group. However, these DNA methylation changes were not correlated with any decreased expression of the corresponding genes. Except for DMRs, the large DMBs between were the two groups were also detected, and DMBs were overlapped with non-PMD or PMD sparse regions on chromosomes which again proved the stability of PMDs through early gestation. The functions of these DMBs still need to be further investigated.

It is important to note that not all DNA methylation changes at promoters and TSS regions were negatively correlated with gene expression. In this particular study, DNA methylation of DMRs mainly increased across gestation which is consistent with the literature that shows overall increases of placental DNA methylation during development after implantation [52-54]. In the present study, only increased DNA methylation in DMRs that overlapped with promoters of *CCR7*, *SLC2A3*, *IL17D* and *EPB42* was correlated with decreased gene expression. CCR7 is found in both syncytiotrophoblasts and villous cytotrophoblasts [55]. It is the receptor for CCL21 chemokine that has been shown to increase the migration of extravillous cytotrophoblasts [56]. The regulation of CCR7 we see could associate with the migration of trophoblasts into the decidua.

It has been suggested that the DMR at *SLC2A3* promoter regions could be important for placental development, since SLC2A3 is a glucose transporter that play a role in glucose supply during early fetal development and reduced expression of *SLC2A3* may related with Intrauterine growth retardation [57, 58]. *EPB42* is an erythrocyte-specific gene and it is involved in iron metabolism. *EPB42* was reported to be

upregulated in peripheral blood mononuclear cells in cancer patients with immune responses to cancer vaccines [59]. However, the function of EPB42 in placenta is not clear. *IL17D* is a cytokine that can regulate cytokine production in endothelial cells [60]. Expression of *IL17D* is decreased in metastatic human tumours [61]. Using samples from first trimester, a previous study has shown that DNA methylation of *IL17D* is significantly lower in placenta than in the corresponding maternal blood [62]. The increase of placental DNA methylation in *IL17D* across early gestation is novel and deserves further investigation.

It has been shown that DNA methylation sites can be used to predict age for numerous human tissues [24]. In most of the studies, DNA methylation sites shared between 450k and EPIC arrays were used for predicting placental gestational age [63, 64] and DNA methylation age is altered in placenta from complicated pregnancies. In this study, we established an epigenetic clock for placenta samples from terminations at 6-23 weeks' gestation using sites from EPIC arrays since a previous study has shown that the accuracy of age estimation does not change between array platforms [65]. To limit the overfitting of the models, we reduced the correlated sites using recursive feature elimination (RFE) method in random forest [66]. We also increased the numbers for the training data set, used 10-fold cross validation to penalize the parameters and validated predictors in a separate test data set compared to a previously published study [67]. More placenta samples will be used in the future to test the gestational age clock and use it to predict DNA methylation age in placentas from healthy and complicated pregnancies.

This study provides the first profiling of DNA methylation of human placenta chorionic villous tissue across early to mid-gestation using Illumina EPIC BeadChip, which interrogates a much larger number of methylation sites. Moreover, RNA-seq assessment of gene expression in the same samples enabled analysis of correlations between differentially methylated genes and gene expression. The limitation of this study is that the cell composition of each placenta sample is not considered in the linear models when detecting differential methylation. An additional method may need to be considered for both DNA methylation and gene expression data to account for changes in chorionic villous tissue composition over time. However, we have considered variation between samples by weighting each sample in the model for differential methylation analyses. A major strength of this study is the large number of placentas sampled and the inclusion of samples up to 23 weeks' gestation.

4.5 Conclusion

This study identified DNA methylation changes during early placental development and improved our understanding of the DNA methylation profile and how it associated with gene expression of placenta. Placental DNA methylation changed throughout 6-23 weeks' gestation and was also apparent from before and after 10 weeks' gestation. Presumably these changes coincide with initiation of maternal blood flow into the intervillous space and rising oxygen tension. Identification and further analyses of associated coding and non-coding gene expression will provide novel insights into early placental development in health and disease.

4.6 Supplementary figures and tables

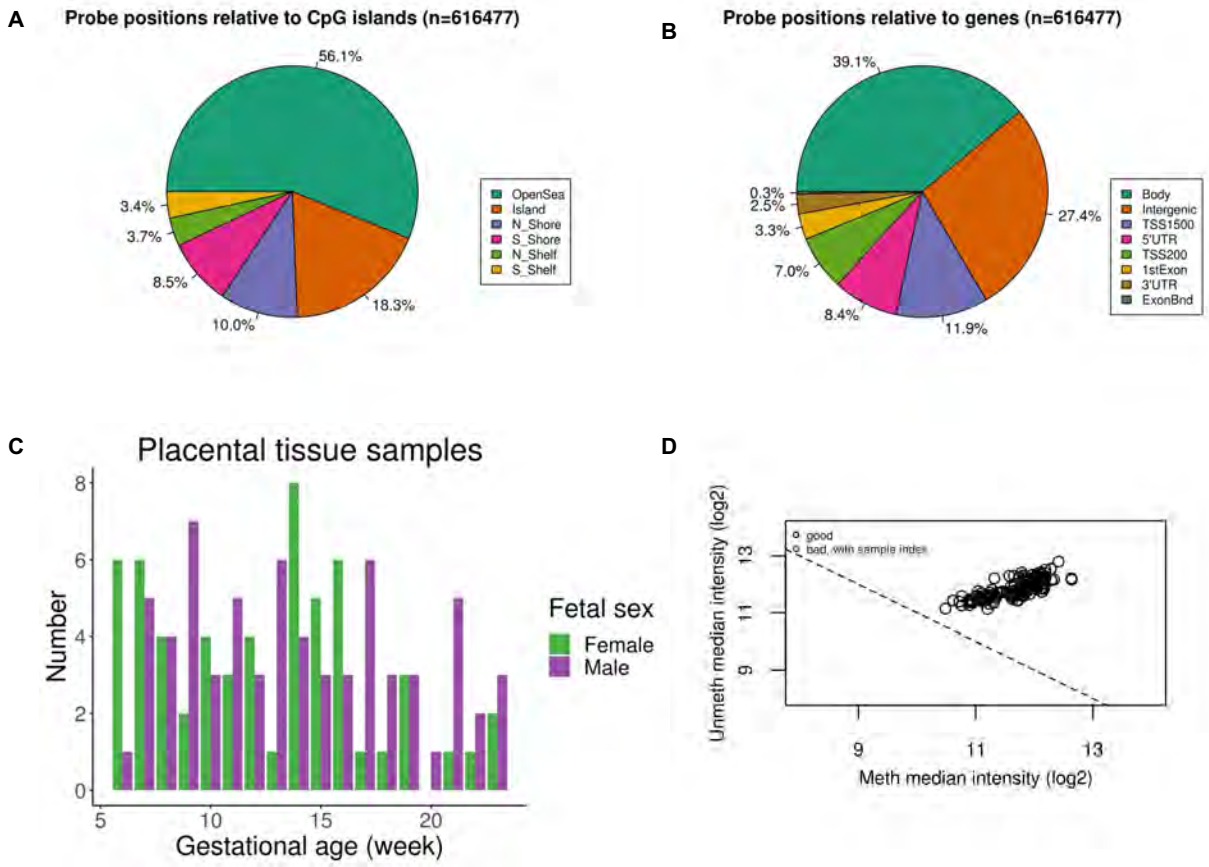

Figure S4.1 Overview of placenta samples and probes used for this study. (A) Positions (relative to CpG islands) of 616766 probes after filtering of failed probes, probes with SNPs and probes on sex chromosomes. (B) Positions of filtered probes relative to gene. (C) Overview of 125 human placental tissue samples (68 males, 63 females) spanning 6 to 23 weeks' gestation profiled in this study. (D) Quality control of all samples showing all samples are of good quality.

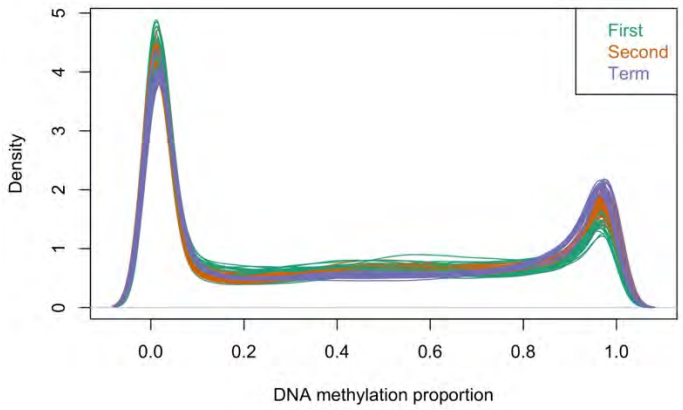

Figure S4.2 Probe density of 319201 common probes between array platforms for all 147 placenta samples. First trimester (n=57), second trimester (n=68), term (n=22).

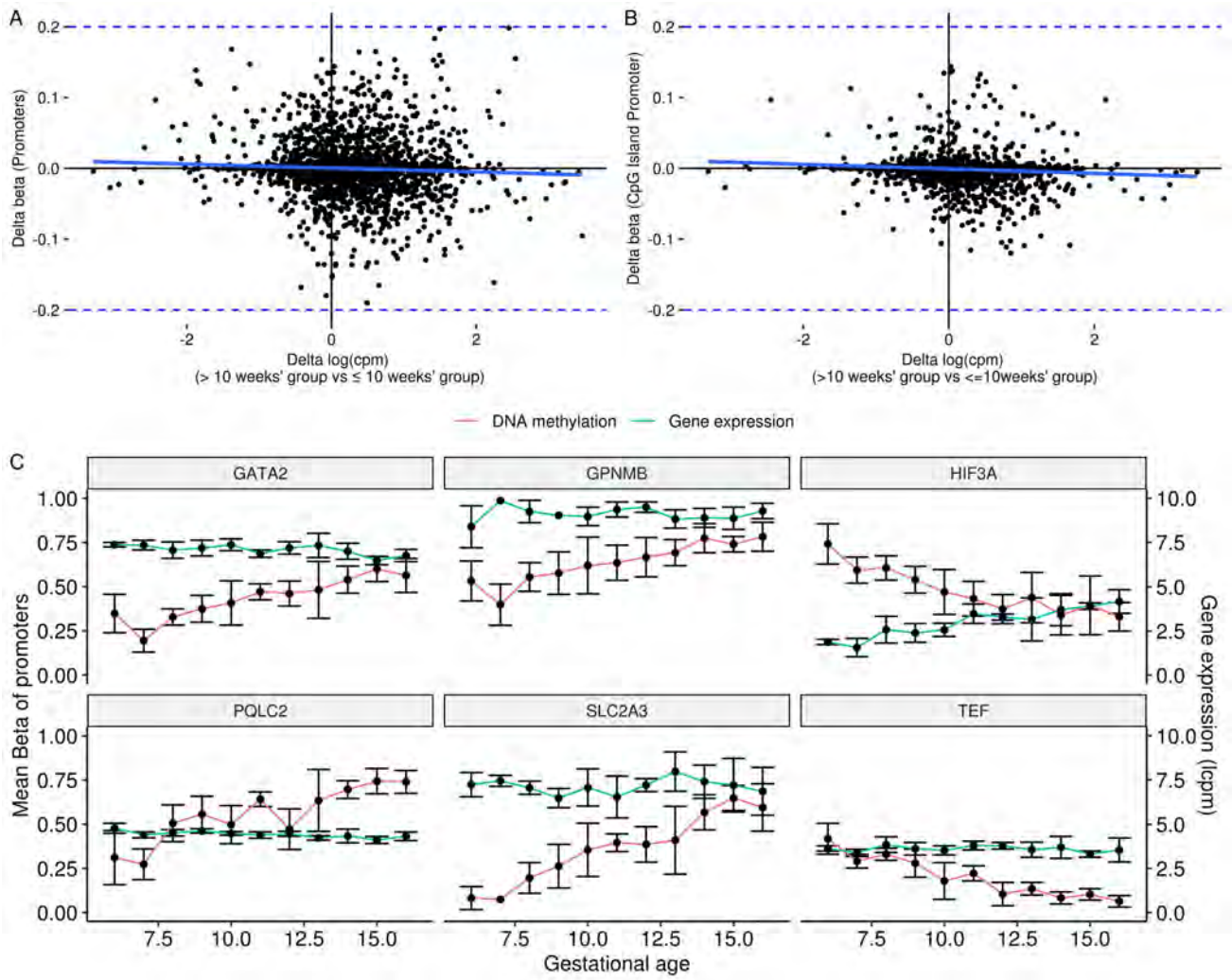

Figure S4.3 Increases of DNA methylation at enhancers and promoter of genes were not significantly correlated with repression of genes in placenta. (A) Gene expression and DNA methylation changes at promoters of each gene in placenta. (B) Gene expression and DNA methylation changes at enhancers of each gene in placenta. (C) DNA methylation and expression of genes with absolute delta beta >0.2 in promoter regions.

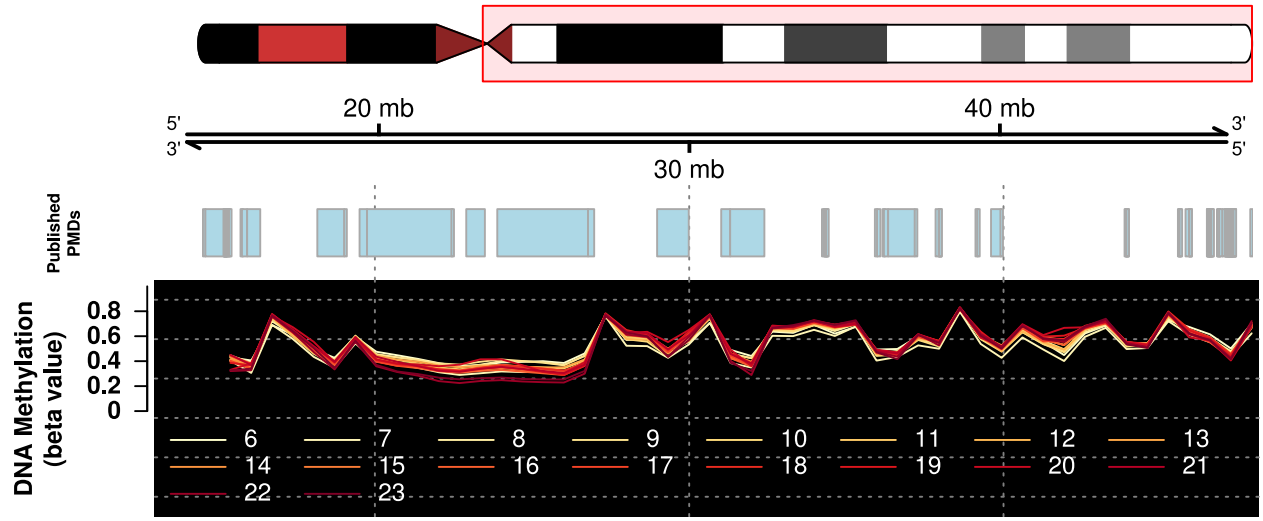

Figure S4.4 The DNA methylation of PMDs is stable across gestation. Representative example on long arm of chromosome 21 of PMDs for all samples from 6 to 23 weeks' gestation.

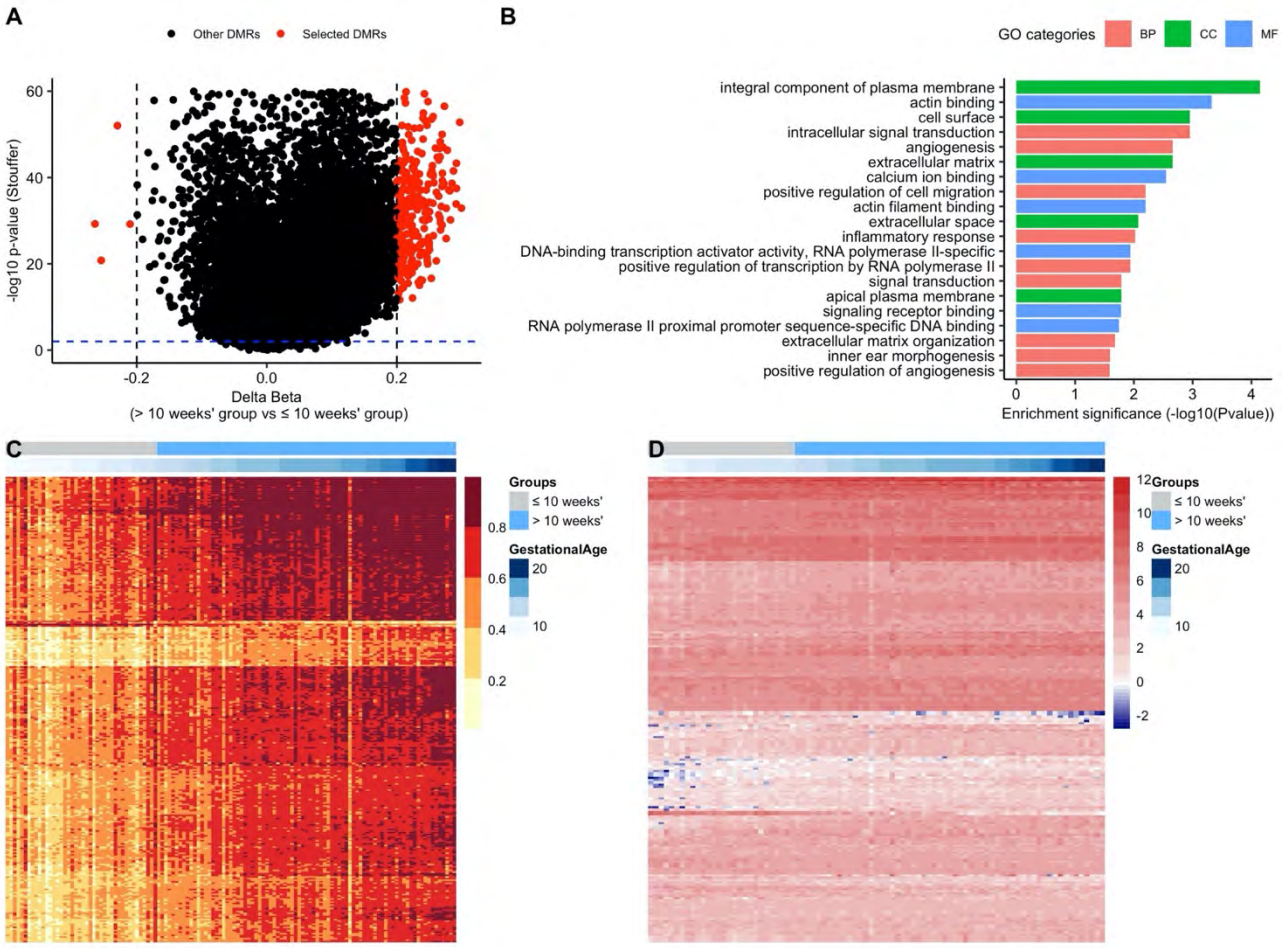

Figure S4.5 DNA methylation changed across gestation (≤ 10 and > 10 weeks' gestation). DMRs identified using all 616766 probes. (A) Volcano plot of all regions identified, threshold of differentially methylated regions (DMRs) were P-value < 0.05 and absolute ΔBeta >0.2. (B) Gene otology of DMRs overlapped genes. (C) DNA methylation of 295 DMRs with P-value < 0.05 and absolute ΔBeta >0.2. (D) Gene expression of DMR overlapped genes, genes with low expression were filtered out.

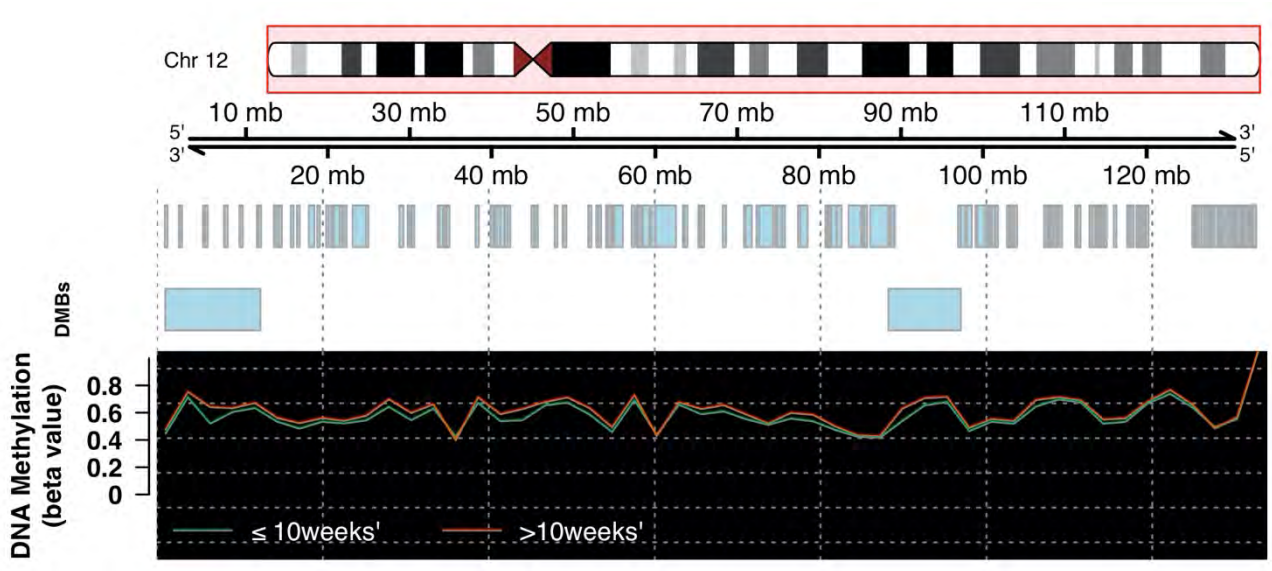

Figure S4.6 DNA methylation changed at DMBs (≤ 10 and > 10 weeks' gestation). For example, there were 2 DMBs identified on chromosome 12, and they are located in non-PMDs or PMD sparse regions.

Table S4.1 Meta data of 131 placenta samples.

Sample ID	Array Date Received	Gestation (weeks)	Trimester	Fetal sex	Outliers
PAC0006	23/3/2017	9	First	F	Outlier
PAC0007	23/3/2017	10	First	M	Standard
PAC0008	23/3/2017	9	First	F	Outlier
PAC0009	23/3/2017	8	First	F	Standard
PAC0010	23/3/2017	8	First	F	Standard
PAC0011	23/3/2017	11	First	M	Standard
PAC0012	23/3/2017	7	First	F	Standard
PAC0013	23/3/2017	10	First	F	Standard
PAC0014	23/3/2017	9	First	M	Standard
PAC0015	23/3/2017	11	First	F	Standard
PAC0016	23/3/2017	9	First	M	Standard
PAC0017	23/3/2017	14	Second	F	Standard
PAC0018	23/3/2017	21	Second	M	Standard
PAC0020	23/3/2017	20	Second	M	Standard
PAC0021	23/3/2017	23	Second	M	Standard
PAC0022	23/3/2017	19	Second	M	Standard
PAC0023	23/3/2017	22	Second	M	Standard
PAC0024	23/3/2017	6	First	F	Standard
PAC0025	23/3/2017	7	First	M	Standard
PAC0026	23/3/2017	16	Second	M	Standard
PAC0027	23/3/2017	16	Second	F	Standard
PAC0029	23/3/2017	17	Second	M	Standard
PAC0030	23/3/2017	13	Second	M	Standard

PAC0031	23/3/2017	16	Second	M	Standard
PAC0032	23/3/2017	13	Second	M	Standard
PAC0033	23/3/2017	22	Second	F	Standard
PAC0034	23/3/2017	8	First	F	Standard
PAC0035	23/3/2017	6	First	F	Outlier
PAC0036	23/3/2017	6	First	F	Outlier
PAC0037	23/3/2017	7	First	F	Standard
PAC0038	23/3/2017	9	First	F	Standard
PAC0039	23/3/2017	6	First	F	Outlier
PAC0040	23/3/2017	13	Second	M	Standard
PAC0041	23/3/2017	6	First	F	Standard
PAC0042	23/3/2017	8	First	M	Standard
PAC0043	23/3/2017	12	First	M	Standard
PAC0044	23/3/2017	9	First	M	Standard
PAC0045	23/3/2017	7	First	M	Standard
PAC0046	23/3/2017	8	First	F	Standard
PAC0047	23/3/2017	10	First	M	Standard
PAC0048	23/3/2017	6	First	M	Standard
PAC0049	23/3/2017	11	First	M	Standard
PAC0050	23/3/2017	9	First	M	Standard
PAC0051	23/3/2017	7	First	F	Standard
PAC0052	23/3/2017	9	First	M	Standard
PAC0053	23/3/2017	16	Second	F	Standard
PAC0054	23/3/2017	14	Second	M	Standard
PAC0055	23/3/2017	6	First	F	Standard
PAC0056	23/3/2017	6	First	F	Standard
PAC0057	1/8/2017	17	Second	M	Standard
PAC0058	23/3/2017	16	Second	F	Standard
PAC0059	23/3/2017	10	First	F	Standard
PAC0060	23/3/2017	9	First	M	Standard
PAC0062	23/3/2017	10	First	M	Standard
PAC0063	23/3/2017	9	First	M	Standard
PAC0064	23/3/2017	7	First	F	Standard
PAC0065	1/8/2017	12	First	M	Standard
PAC0069	1/8/2017	15	Second	M	Standard
PAC0070	1/8/2017	13	Second	M	Standard
PAC0071	1/8/2017	12	First	F	Standard
PAC0072	1/8/2017	12	First	M	Standard
PAC0074	1/8/2017	13	Second	F	Standard
PAC0075	1/8/2017	23	Second	M	Standard
PAC0076	1/8/2017	18	Second	M	Standard
PAC0077	1/8/2017	19	Second	M	Standard
PAC0078	1/8/2017	14	Second	F	Standard

PAC0083	1/8/2017	8	First	M	Standard
PAC0084	1/8/2017	19	Second	F	Standard
PAC0086	1/8/2017	11	First	M	Standard
PAC0087	1/8/2017	14	Second	M	Standard
PAC0088	1/8/2017	11	First	F	Standard
PAC0091	1/8/2017	11	First	M	Standard
PAC0093	1/8/2017	15	Second	M	Standard
PAC0097	1/8/2017	16	Second	F	Standard
PAC0098	1/8/2017	14	Second	M	Standard
PAC0099	1/8/2017	14	Second	F	Standard
PAC0100	1/8/2017	15	Second	F	Standard
PAC0102	1/8/2017	15	Second	F	Standard
PAC0103	1/8/2017	15	Second	M	Standard
PAC0105	1/8/2017	12	First	F	Standard
PAC0107	1/8/2017	10	First	F	Standard
PAC0108	1/8/2017	18	Second	M	Standard
PAC0109	1/8/2017	14	Second	F	Standard
PAC0111	1/8/2017	16	Second	F	Standard
PAC0114	1/8/2017	21	Second	M	Standard
PAC0117	1/8/2017	21	Second	M	Standard
PAC0118	1/8/2017	23	Second	F	Standard
PAC0120	1/8/2017	14	Second	M	Standard
PAC0122	1/8/2017	10	First	F	Standard
PAC0124	1/8/2017	17	Second	M	Standard
PAC0127	1/8/2017	18	Second	F	Standard
PAC0129	1/8/2017	14	Second	F	Standard
PAC0131	1/8/2017	18	Second	M	Standard
PAC0134	1/8/2017	16	Second	F	Standard
PAC0139	1/8/2017	22	Second	M	Standard
PAC0140	1/8/2017	21	Second	M	Standard
PAC_0019	1/2/19	17	Second	M	Standard
PAC_0068	1/2/19	13	Second	M	Standard
PAC_0092	1/2/19	13	Second	M	Standard
PAC_0094	1/2/19	11	First	M	Standard
PAC_0101	1/2/19	14	Second	F	Standard
PAC_0106	1/2/19	11	First	F	Standard
PAC_0110	1/2/19	15	Second	F	Standard
PAC_0113	1/2/19	12	First	F	Standard
PAC_0121	1/2/19	14	Second	F	Standard
PAC_0125	1/2/19	15	Second	F	Standard
PAC_0126	1/2/19	12	First	F	Standard
PAC_0128	1/2/19	14	Second	F	Standard
PAC_0130	1/2/19	17	Second	M	Standard

PAC_0135	1/2/19	16	Second	M	Standard
PAC_0137	1/2/19	17	Second	M	Standard
PAC_0141	1/2/19	8	First	M	Standard
PAC_0148	1/2/19	8	First	M	Standard
PAC_0183	1/2/19	15	Second	F	Standard
PAC_0186	1/2/19	7	First	M	Standard
PAC_0191	1/2/19	19	Second	F	Standard
PAC_0192	1/2/19	21	Second	M	Standard
PAC_0193	1/2/19	6	First	F	Standard
PAC_0196	1/2/19	21	Second	F	Standard
PAC_0198	1/2/19	7	First	M	Standard
PAC_0200	1/2/19	7	First	F	Standard
PAC_0202	1/2/19	6	First	M	Outlier
PAC_0203	1/2/19	19	Second	F	Standard
PAC_0204	1/2/19	7	First	M	Standard
PAC_0208	1/2/19	7	First	F	Standard
PAC_0211	1/2/19	17	Second	F	Standard
PAC_0214	1/2/19	9	First	F	Standard
PAC_0215	1/2/19	23	Second	M	Standard
PAC_0219	1/2/19	6	First	F	Standard
PAC_0221	1/2/19	19	Second	M	Standard
PAC_0222	1/2/19	23	Second	F	Standard

Table S4.2 Number of regions and probes in PMDs.

Regions	Number of regions in PMDs	Number of regions not in PMDs
Gene	9684	13372
Gene body	3472 (44395 probes)	16590 (230669 probes)
Promoter	1118 (1253 probes)	13795 (17587 probes)
Promoter in placenta	460 (547 probes)	10140 (13526 probes)
Placenta-specific promoter	2 (2 probes)	4 (4 probes)
CpG island	31 (52 probes)	23483 (112229 probes)
CpG island in gene body	10 (20 probes)	12358 (41270 probes)
CpG island in promoter	1 (1 probes)	6273 (11852 probes)
Enhancer	2742 (2977 probes)	15596 (17756 probes)
Enhancer in placenta	275 (304 probes)	3392 (4209 probes)
Placenta-specific enhancer	1 (1 probes)	3 (4 probes)

Table S4.3 Gene ontology for genes in PMDs.

ID	Gene Ratio	Background Ratio	P value	Adjust P value	Count	Cat	Terms
GO:0050911	298/1367	427/17653	2.92E-237	1.25E-233	298	BP	detection of chemical stimulus involved in sensory perception of smell
GO:0007608	301/1367	453/17653	1.24E-229	2.66E-226	301	BP	sensory perception of smell
GO:0031424	97/1367	226/17653	2.67E-48	2.85E-45	97	BP	keratinization
GO:0008544	115/1367	462/17653	2.03E-30	1.09E-27	115	BP	epidermis development
GO:0018149	37/1367	59/17653	8.26E-27	3.92E-24	37	BP	peptide cross-linking
GO:0042742	51/1367	293/17653	3.53E-08	1.26E-05	51	BP	defence response to bacterium
GO:0019730	27/1367	108/17653	3.54E-08	1.26E-05	27	BP	antimicrobial humoral response
GO:0061844	17/1367	59/17653	1.42E-06	4.35E-04	17	BP	antimicrobial humoral immune response mediated by antimicrobial peptide
GO:0006959	51/1367	329/17653	1.43E-06	4.35E-04	51	BP	humoral immune response
GO:1900003	8/1367	15/17653	4.98E-06	0.00127478	8	BP	regulation of serine-type endopeptidase activity
GO:1902571	8/1367	15/17653	4.98E-06	0.00127478	8	BP	regulation of serine-type peptidase activity
GO:0007210	10/1367	24/17653	5.37E-06	0.00127478	10	BP	serotonin receptor signalling pathway
GO:0098664	10/1367	24/17653	5.37E-06	0.00127478	10	BP	G-protein coupled serotonin receptor signalling pathway
GO:0072677	11/1367	30/17653	7.94E-06	0.00178553	11	BP	eosinophil migration
GO:0048245	10/1367	26/17653	1.26E-05	0.00269157	10	BP	eosinophil chemotaxis
GO:0004984	298/1305	427/17548	4.38E-243	3.34E-240	298	MF	olfactory receptor activity
GO:0005549	72/1305	95/17548	1.07E-61	4.08E-59	72	MF	odorant binding
GO:0004867	25/1305	95/17548	1.69E-08	3.22E-06	25	MF	serine-type endopeptidase inhibitor activity
GO:1903231	35/1305	196/17548	1.05E-06	9.98E-05	35	MF	mRNA binding involved in posttranscriptional gene silencing
GO:0019865	10/1305	22/17548	1.41E-06	1.19E-04	10	MF	immunoglobulin binding
GO:0022824	16/1305	57/17548	2.49E-06	1.73E-04	16	MF	transmitter-gated ion channel activity
GO:0022835	16/1305	57/17548	2.49E-06	1.73E-04	16	MF	transmitter-gated channel activity
GO:0004497	22/1305	101/17548	3.90E-06	2.12E-04	22	MF	monooxygenase activity
GO:0030280	8/1305	17/17548	1.21E-05	4.63E-04	8	MF	structural constituent of epidermis
GO:0098960	12/1305	40/17548	2.14E-05	7.77E-04	12	MF	postsynaptic neurotransmitter receptor activity
GO:0015347	10/1305	33/17548	9.48E-05	0.00288675	10	MF	sodium-independent organic anion transmembrane transporter activity

GO:0004252	35/1305	248/17548	1.85E-04	0.0052254	35	MF	serine-type endopeptidase activity
GO:0099095	7/1305	19/17548	2.81E-04	0.0072781	7	MF	ligand-gated anion channel activity
GO:0019825	11/1305	45/17548	3.55E-04	0.00817538	11	MF	oxygen binding
GO:0004806	8/1305	26/17548	4.25E-04	0.00952196	8	MF	triglyceride lipase activity
GO:0001533	37/1457	65/18698	1.46E-24	6.58E-22	37	CC	cornified envelope
GO:0005882	62/1457	198/18698	2.90E-22	6.53E-20	62	CC	intermediate filament
GO:0045111	62/1457	237/18698	7.07E-18	1.06E-15	62	CC	intermediate filament cytoskeleton
GO:0034702	44/1457	296/18698	2.53E-05	0.00228637	44	CC	ion channel complex
GO:1990351	46/1457	329/18698	7.81E-05	0.0044024	46	CC	transporter complex

Table S4.4 DMRs identified using different subsets of DNA methylation sites.

Sites used for DMR analyses	Chr	start	end	width	Number of sites in DMR	Overlapping genes	Stouffer	Mean difference
use all sites	chr17	56606016	56607967	1952	8	SNORA69, RP11-112H10.4, SEPT4	3.19E-113	0.19394788
use all sites	chr1	84628924	84630556	1633	11	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, PRKACB, SNORA2, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	2.26E-97	0.2332764
use all sites	chr22	24914326	24916084	1759	6	UPB1, AP000355.2	1.88E-94	0.23008898
use all sites	chr5	149555647	149558804	3158	7	CDX1, SNORA68, SNORA57, SNORD45, SNORD95	3.50E-89	0.19471433
use all sites	chr1	202171585	202172912	1328	8	snoU13, Y_RNA, SNORD112, U3, SNORA51, SNORA25, SCARNA20, U6, SNORA70, LGR6, SNORA77, SNORA26, SNORA72, U8, SNORD60, SNORD116	5.66E-84	0.19725395
use all sites	chr8	127516332	127516878	547	4	NA	7.49E-79	0.29377881
use all sites	chr5	132589269	132592155	2887	8	SNORA27, SNORA68, SNORA57, 7SK, FSTL4, SNORD45, SNORD95	6.03E-74	0.18290804
use all sites	chr5	75903397	75903700	304	6	IQGAP2, snoZ6, SNORA27, SNORA68, RPS23P5, SNORA57, 7SK, CTD-2236F14.1, SNORD45	5.22E-74	0.26674486
use all sites	chr1	10719619	10720993	1375	9	snoU13, Y_RNA, SCARNA16, SCARNA21, SNORD112, SNORA62, SNORA2, U3, SNORA51, SCARNA20, SNORA67, SNORA70, SNORA77, SNORA26, U8, CASZ1, SCARNA11, SNORA31, SNORA42, SNORA40, snoU109, SNORD60	1.28E-71	0.18655665
use all sites	chr1	236004282	236005962	1681	5	snoU13, Y_RNA, SNORD112, SNORA25, LYST	2.69E-70	0.21678288
use all sites	chr15	75039788	75041386	1599	7	CYP1A2	2.75E-63	0.20616141
use all sites	chr16	75096111	75097612	1502	6	ZNRF1, SNORD33	8.02E-68	0.21122186
use all sites	chr1	215178601	215179216	616	8	snoU13, Y_RNA, SNORD112, U3, SNORA51, SNORA25, SNORA70, KCNK2, SNORA72, U8, SNORD116	4.25E-67	0.21997604
use all sites	chr3	46785957	46787073	1117	6	U7, SNORD77, SNORA33, SNORA81, SNORD5, PRSS50, PRSS45, SNORD38, SNORD63, Metazoa_SRP, SNORA18	1.37E-67	0.20030576
use all sites	chr2	25551575	25552751	1177	6	SNORA73, SNORA64, SNORA74, snR65, SCARNA6, SNORD39, SNORD18, DNMT3A, MIR1301, SNORA75, SNORA48, SNORD56, SNORA43, SNORA1	1.37E-60	0.19986964

use all sites	chr3	108180778	108181837	1060	5	U7, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, SNORD61, MYH15, SNORA24, Metazoa_SRP, SNORA18	3.14E-65	0.24076955
use all sites	chr8	130695235	130696901	1667	7	RP11-419K12.1	7.54E-65	0.23573812
use all sites	chr8	130691264	130693960	2697	7	CCDC26, RP11-419K12.1	3.63E-59	0.19679063
use all sites	chr13	46757395	46757637	243	4	SNORD36, LCP1, SNORD37	1.33E-61	0.2232919
use all sites	chr15	69605276	69606334	1059	5	PAQR5	4.06E-60	0.23292118
use all sites	chr1	84863547	84864880	1334	7	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, SNORA2, DNASE2B, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	5.26E-58	0.19742397
use all sites	chr2	71787501	71787827	327	5	SNORA73, SNORA64, SNORD75, SNORA12, SNORA74, DYSF, SNORA19, snR65, 5S_rRNA, SNORA4, SNORD11, SNORA41, SCARNA6, SNORD39, SNORD18, SNORA36, SNORA75, ACA59, SNORA48, SNORD56, SNORA43, SNORA1, Vault	2.59E-58	0.23699672
use all sites	chr6	2750213	2752360	2148	6	MYLK4	1.78E-54	0.22368009
use all sites	chr2	60618524	60619220	697	6	SNORA73, SNORA64, SNORD75, SNORA12, SNORA74, snR65, 5S_rRNA, SNORA4, SNORD11, SNORA41, SCARNA6, SNORD39, SNORD18, SNORA36, SNORA75, SNORA48, SNORD56, SNORA43, SNORA1, Vault	4.47E-46	0.19439776
use all sites	chr10	116417809	116419517	1709	5	ABLIM1	3.11E-57	0.26260368
use all sites	chr12	105322376	105323744	1369	5	SNORA9, SLC41A2	8.33E-53	0.19969042
use all sites	chr12	79811775	79813192	1418	5	SYT1, MIR1252, SNORA9, RP1-78O14.1, snoMe28S-Am2634	7.53E-55	0.24575267
use all sites	chr2	110883862	110885303	1442	5	SNORA64, SNORA12, SNORA74, SNORA19, LIMS3, snR65, 5S_rRNA, SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, ACA59, NPHP1, AC013268.1, SNORA48, SNORD56, SNORA43, SNORA1, Vault	5.30E-49	0.25078368
use all sites	chr16	67421302	67422624	1323	4	SNORD111, SNORD33	1.17E-55	0.2276662
use all sites	chr7	107331385	107332802	1418	5	SLC26A4	3.07E-54	0.23108318
use all sites	chr5	143301646	143303264	1619	4	SNORA68, SNORA57, 7SK, SNORD45, SNORD95	1.76E-52	0.21077111
use all sites	chr12	5673852	5675505	1654	6	ANO2	3.96E-52	0.21990575
use all sites	chr5	95321963	95323104	1142	5	CTD-2337A12.1, snoZ6, SNORA27, SNORA68, SNORA57, 7SK, AC008592.7, SNORD45, SNORD95	9.42E-48	0.25835775
use all sites	chr20	23011983	23012907	925	4	RP4-753D10.3	6.74E-54	0.22611532
use all sites	chr4	147032361	147033202	842	4	LINC01095, SNORD65	7.67E-53	0.24572518
use all sites	chr2	111951702	111953038	1337	4	SNORA64, SNORA12, SNORA74, SNORA19, snR65, 5S_rRNA, SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, ACA59, SNORA48, SNORD56, SNORA43, SNORA1, Vault	1.63E-53	0.28777125
use all sites	chr22	20923127	20923814	688	3	MED15	9.55E-53	-0.2224359
use all sites	chr6	6622864	6624448	1585	5	LY86, LY86-AS1, SNORA20	1.11E-50	0.20007951
use all sites	chr11	27602211	27603216	1006	6	BDNF-AS	3.27E-52	0.23137716
use all sites	chr6	76108836	76109818	983	4	SNORA38, SNORA8, SCARNA15, SNORD28, FILIP1, SNORA20	1.06E-50	0.26031588
use all sites	chr7	93706476	93707637	1162	5	NA	1.30E-49	0.22284779
use all sites	chr15	45802899	45803341	443	4	RP11-519G16.3, HMGN2P46, SLC30A4	2.68E-50	0.33344958

use all sites	chr6	76073316	76074059	744	3	SNORA38, SNORA8, SCARNA15, RP11-415D17.3, SNORD28, FILIP1, SNORA20	6.64E-49	0.31724684
use all sites	chr11	105010284	105011637	1354	6	CARD18	3.87E-40	0.18894314
use all sites	chr11	74424489	74424521	33	3	CHRDL2, SNORD43	7.44E-49	0.19845987
use all sites	chr10	126826304	126827067	764	5	CTBP2	2.29E-49	0.2089085
use all sites	chr1	120453393	120454387	995	4	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, SNORA42, SNORA40, NOTCH2, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	1.13E-47	0.28163339
use all sites	chr3	36902435	36903038	604	5	U7, SNORD77, SNORA33, SNORA81, SNORD5, TRANK1, SNORD38, SNORD63, Metazoa_SRP, SNORA18	1.99E-45	0.25504301
use all sites	chr12	6553092	6554051	960	5	CD27, CD27-AS1	7.68E-48	0.19256734
use all sites	chr1	84257836	84259317	1482	3	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, RP5-836J3.1, SNORA2, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, RP11-475O6.1, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	2.88E-46	0.22738067
use all sites	chr17	3791228	3792074	847	4	CAMKK1	1.58E-47	0.22873749
use all sites	chr18	56266818	56268032	1215	4	ALPK2	1.86E-47	0.24816158
use all sites	chr4	15230035	15230432	398	3	SNORA3, RP11-665G4.1	4.93E-47	0.29887876
use all sites	chr3	37795590	37796300	711	3	U7, ITGA9, SNORD77, SNORA33, SNORA81, SNORD5, AC093415.2, SNORD38, SNORD63, Metazoa_SRP, SNORA18	3.83E-46	0.33459445
use all sites	chr5	146298512	146300238	1727	4	SNORA68, SNORA57, SNORD45, PPP2R2B, PPP2R2B-IT1, SNORD95	4.65E-47	0.19859608
use all sites	chr11	76744970	76745431	462	3	B3GNT6	9.11E-47	0.22831784
use all sites	chr3	157213371	157213897	527	4	SNORA81, SNORD66, SNORD2, VEPH1, SNORA18, U4	4.79E-47	0.269309
use all sites	chr2	208037561	208037968	408	3	SCARNA6, SNORD39, SNORA75, SNORA48	1.41E-46	0.29297356
use all sites	chr12	7241449	7242709	1261	4	C1R, C1RL	1.25E-44	0.20024575
use all sites	chr10	89365807	89367384	1578	4	SNORA17	7.65E-45	0.20921993
use all sites	chr1	174668743	174670136	1394	5	snoU13, Y_RNA, SCARNA16, SNORD112, SNORA63, U3, SNORA51, SNORA25, SNORD59, SCARNA20, RABGAP1L, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, snoU109, SNORD60, SNORD116	4.41E-36	0.21781028
use all sites	chr1	55951958	55952111	154	3	snoU13, Y_RNA, SCARNA16, U1, SCARNA18, SCARNA24, SNORD112, SNORA62, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, SNORA58, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, snoU109, SNORD60, SNORD116	4.65E-45	0.21468964
use all sites	chr7	107796618	107798547	1930	5	NRCAM	4.30E-44	0.19746075
use all sites	chr7	93189633	93190462	830	5	CALCR	1.81E-40	0.21181492
use all sites	chr10	60577367	60579479	2113	4	BICC1, SNORA71	1.49E-45	0.25042561
use all sites	chr1	110761524	110763078	1555	3	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, KCNC4, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	7.37E-45	0.1970481
use all sites	chr1	206226009	206227162	1154	3	snoU13, Y_RNA, SNORD112, U3, SNORA51, SNORA25, SNORA70, AVPR1B, SNORA26, SNORA72, U8, SNORD60, SNORD116	9.16E-45	0.2436713
use all sites	chr10	74652084	74653449	1366	4	OIT3, SNORA71, SNORA17	3.96E-43	0.20192562

use all sites	chr15	39541977	39543043	1067	4	C15orf54, RP11-624L4.1	2.29E-42	0.24079396
use all sites	chr5	60932581	60933135	555	3	SNORA50, snoZ6, SNORA27, SNORA68, RPS23P5, SNORA57, SNORA76, 7SK, SNORD45	6.22E-44	0.22263006
use all sites	chr4	147029689	147030199	511	3	SNORD65	5.13E-44	0.26601824
use all sites	chr1	160342957	160344469	1513	4	snoU13, Y_RNA, SCARNA16, SNORD112, SNORA63, U3, SNORA51, SNORA25, SNORD59, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, SNORA40, ACA64, SNORD78, snoU109, SNORD60, SNORD116	2.16E-44	0.21276169
use all sites	chr5	75671504	75672341	838	3	snoZ6, SNORA27, SNORA68, RPS23P5, SNORA57, 7SK, SNORD45	1.75E-43	0.25131944
use all sites	chr17	38716802	38718229	1428	4	SNORA69, CCR7	8.38E-33	0.2388912
use all sites	chr3	185438794	185439589	796	3	SNORA81, C3orf65, SNORD2, IGF2BP2, U4	1.45E-42	0.19828005
use all sites	chr2	48950153	48951719	1567	4	SNORA73, SNORA64, SNORD75, STON1-GTF2A1L, GTF2A1L, SNORA74, snR65, 5S_rRNA, SNORD11, SNORA41, SCARNA6, SNORD39, SNORD18, SNORA36, LHCGR, SNORA75, SNORA48, SNORD56, SNORA43, SNORA1, Vault	3.04E-33	0.22659371
use all sites	chr3	175493926	175495361	1436	4	NAALADL2, SNORA81, SNORD66, SNORD2, NAALADL2-AS1, SNORA18, U4	1.28E-43	0.22384743
use all sites	chr2	27199491	27201619	2129	4	SNORA73, MAPRE3, SNORA64, SNORA74, snR65, SCARNA6, SNORD39, SNORD18, AC013472.4, SNORA75, SNORA48, SNORD56, SNORA43, SNORA1, Vault	2.05E-35	0.21034855
use all sites	chr5	43044573	43045001	429	5	AC025171.1, SNORA27, SNORA68, RPS23P5, SNORA57, CTD-2201E18.3, SNORA76, 7SK, SNORD45	1.74E-43	0.17869904
use all sites	chr6	143192478	143193920	1443	4	SNORD28, HIVEP2, SNORA20	2.69E-40	0.19530545
use all sites	chr18	10475860	10477185	1326	3	APCDD1	2.41E-42	0.18573901
use all sites	chr18	60862198	60863116	919	3	BCL2	6.99E-42	0.21445041
use all sites	chr12	57405251	57406900	1650	3	TAC3, snoMe28S-Am2634	1.51E-42	0.21675969
use all sites	chr2	111871839	111873850	2012	6	SNORA64, SNORA12, SNORA74, SNORA19, ACOXL, snR65, 5S_rRNA, SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, ACA59, AC096670.3, SNORA48, SNORD56, SNORA43, SNORA1, Vault	2.07E-38	0.22397651
use all sites	chr7	139433877	139434142	266	3	HIPK2	3.08E-42	0.27671406
use all sites	chr10	87364316	87365005	690	5	SNORA17, GRID1	1.64E-41	0.25385484
use all sites	chr7	139408541	139409421	881	3	HIPK2	1.94E-41	0.2446231
use all sites	chr16	57033268	57034444	1177	3	NLRC5, SNORD111, SNORD33	1.38E-41	0.1946079
use all sites	chr6	36804444	36805384	941	3	SNORA38, CPNE5, SNORA20	1.60E-41	0.21524224
use all sites	chr5	179494929	179495386	458	3	RNF130, SNORD95	1.07E-41	0.25771953
use all sites	chr6	128239611	128241185	1575	5	SNORA8, SNORD28, THEMIS, SNORA20	3.20E-39	0.22171412
use all sites	chr5	111333193	111333693	501	4	NREP-AS1, snoZ6, SNORA27, SNORA68, SNORA57, 7SK, SNORD45, SNORD95	7.02E-42	0.21732233
use all sites	chr2	218879105	218881298	2194	4	SCARNA6, SNORD39, SNORA75, SNORA48	2.33E-35	0.19589115
use all sites	chr12	76413973	76414729	757	3	SNORA9, snoMe28S-Am2634	1.40E-40	0.19959356
use all sites	chr7	115608503	115608801	299	3	TFEC	7.59E-41	0.33238299

use all sites	chr2	69042739	69043808	1070	3	SNORA73, SNORA64, SNORD75, SNORA12, SNORA74, ARHGAP25, SNORA19, snR65, 5S_rRNA, SNORA4, SNORD11, SNORA41, SCARNA6, SNORD39, SNORD18, SNORA36, SNORA75, ACA59, SNORA48, SNORD56, SNORA43, SNORA1, Vault	5.74E-40	0.26944867
use all sites	chr7	158673837	158673979	143	3	WDR60	4.76E-41	0.2477287
use all sites	chr11	115111606	115111974	369	3	CADM1	1.64E-39	0.24192893
use all sites	chr5	90353937	90355362	1426	4	GPR98, snoZ6, SNORA27, SNORA68, SNORA57, 7SK, SNORD45, SNORD95	5.42E-40	0.21999739
use all sites	chr3	171494766	171496417	1652	3	SNORA81, SNORD66, SNORD2, PLD1, SNORA18, U4	1.31E-39	0.20910561
use all sites	chr15	43513306	43513563	258	3	EPB42	2.11E-38	0.19142707
use all sites	chr4	25880545	25882914	2370	5	SMIM20, SNORD74, SNORA3, snoR442, snoU2_19, SNORD65	8.11E-40	0.19672011
use all sites	chr12	48257191	48258214	1024	3	VDR, snoMe28S-Am2634	1.15E-39	0.18869123
use all sites	chr12	50321909	50323001	1093	4	RP11-70F11.8, snoMe28S-Am2634	3.50E-39	0.23679571
use all sites	chr2	9504738	9505705	968	3	SNORA73, ASAP2, SNORA64, snR65, SNORA75, SNORA48, SNORD56, SNORA43	5.94E-39	0.21718589
use all sites	chr18	60848167	60848973	807	3	BCL2	3.12E-40	0.2098908
use all sites	chr7	16308656	16308850	195	4	ISPD-AS1, ISPD	6.80E-40	0.17558463
use all sites	chr10	23363700	23363820	121	3	NA	3.10E-39	0.26329941
use all sites	chr4	170901597	170902837	1241	4	SNORD65	1.68E-39	0.20764897
use all sites	chr22	22383893	22385134	1242	3	PRAMENP	4.73E-39	0.21254684
use all sites	chr4	88719777	88720640	864	6	SNORD74, SNORD50, SNORA11, snoU2_19, SNORD65	1.74E-32	0.17960843
use all sites	chr7	83270795	83271654	860	3	SEMA3E	7.07E-38	0.22981519
use all sites	chr2	152739855	152740431	577	4	5S_rRNA, SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, ACA59, SNORA48, CACNB4, SNORD56, SNORA43, SNORA1, Vault	2.26E-39	0.28018407
use all sites	chr3	137851111	137852643	1533	3	SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, A4GNT, Metazoa_SRP, SNORA18, U4	1.16E-38	0.25631895
use all sites	chr2	171459244	171459744	501	3	MYO3B, SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, ACA59, SNORA48, SNORA43, SNORA1, Vault	1.62E-38	0.21994238
use all sites	chr10	91136405	91137460	1056	4	SNORA17, LIPA	1.46E-34	0.21744576
use all sites	chr9	128604965	128605974	1010	3	PBX3	1.13E-38	0.20404308
use all sites	chr20	47923658	47924116	459	3	NA	1.38E-37	0.19675059
use all sites	chr2	201262081	201262839	759	3	SPATS2L, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, SNORA48, SNORA1, Vault	3.35E-38	0.31216197
use all sites	chr6	143159647	143159766	120	3	SNORD28, HIVEP2, SNORA20	7.52E-38	0.2799306
use all sites	chr14	96127564	96128403	840	3	TCL6	1.03E-35	0.32303142
use all sites	chr7	43134845	43135798	954	3	SNORA22, SNORA15	9.85E-38	0.21631167

use all sites	chr12	94228925	94229919	995	4	CRADD, SNORA9, snoMe28S-Am2634	7.77E-34	0.20551334
use all sites	chr1	113047945	113049100	1156	3	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, WNT2B, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	6.27E-38	0.18825254
use all sites	chr9	131002641	131003043	403	3	DNM1	9.65E-38	0.19236516
use all sites	chr9	16624281	16624542	262	3	SNORD27, BNC2	7.89E-38	0.2459992
use all sites	chr11	78458944	78460645	1702	4	RP11-673F18.1, TENM4	2.19E-36	0.26026115
use all sites	chr1	90253388	90254168	781	3	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, SNORA2, SNORD81, RP11-302M6.4, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	9.04E-38	0.23402128
use all sites	chr11	31789429	31790511	1083	5	ELP4, Z83001.1	6.62E-34	0.1945806
use all sites	chr6	143178039	143179316	1278	3	SNORD28, HIVEP2, SNORA20	2.03E-36	0.18534705
use all sites	chr17	38721675	38721942	268	3	SNORA69, CCR7	1.79E-37	0.21504456
use all sites	chr11	35606942	35607359	418	4	NA	1.84E-36	0.20463823
use all sites	chr4	25028735	25030010	1276	4	SNORD74, SNORA3, snoR442, LGI2, snoU2_19, SNORD65	1.23E-35	0.22019284
use all sites	chr17	17840911	17841862	952	3	SNORA69, TOM1L2	7.07E-35	0.24039485
use all sites	chr8	29686067	29687066	1000	4	SNORA7, RP11-94H18.1	2.46E-35	0.21195064
use all sites	chr1	85870184	85870653	470	3	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, SNORA2, RP11-131L23.1, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, DDAH1, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	5.25E-34	0.2564641
use all sites	chr3	148581801	148583152	1352	4	CPA3, SNORA81, SNORD66, SNORD2, RP11-680B3.2, Metazoa_SRP, SNORA18, U4	1.25E-37	0.26149987
use all sites	chr11	19612894	19613613	720	4	LSP1, NAV2	2.77E-37	0.24218497
use all sites	chr2	211180001	211181383	1383	3	SCARNA6, SNORD39, SNORA75, SNORA48	1.26E-33	0.283936
use all sites	chr3	156840234	156841889	1656	4	SNORA81, SNORD66, SNORD2, LINC00880, SNORA18, U4	1.17E-36	0.19260793
use all sites	chr9	130115605	130116418	814	3	GARNL3	5.33E-36	0.21608444
use all sites	chr3	137492780	137493800	1021	3	RP11-2A4.3, SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, Metazoa_SRP, SNORA18, U4	9.08E-35	0.22796492
use all sites	chr2	63670795	63671845	1051	3	SNORA73, SNORA64, SNORD75, SNORA12, SNORA74, SNORA19, snR65, 5S_rRNA, SNORA4, SNORD11, SNORA41, SCARNA6, SNORD39, SNORD18, SNORA36, SNORA75, WDPCP, SNORA48, SNORD56, SNORA43, SNORA1, Vault	4.19E-35	0.24357739
use all sites	chr6	122531803	122532605	803	3	SNORA8, SNORD28, SNORA20	2.39E-35	0.24041079
use all sites	chr10	123260333	123260551	219	3	FGFR2	8.96E-36	0.36592396
use all sites	chr3	9954087	9954273	187	3	IL17RE, SNORD38	2.01E-35	0.21062065
use all sites	chr11	122797295	122797732	438	3	C11orf63	1.38E-33	0.26638281
use all sites	chr11	16424708	16425891	1184	4	LSP1, snoMBII-202, SOX6	7.19E-36	0.21477755

use all sites	chr17	17635366	17636909	1544	3	SNORA69, RAI1	1.47E-34	0.22103994
use all sites	chr7	8156924	8157910	987	3	ICA1	1.38E-34	0.29063223
use all sites	chr1	45079807	45080600	794	3	snoU13, Y_RNA, SCARNA16, U1, SCARNA17, SCARNA18, SCARNA24, SNORD112, SNORA62, SNORA63, RNF220, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, snoU109, SNORD60, SNORD116	3.72E-34	0.19469353
use all sites	chr21	41029374	41029700	327	4	B3GALT5, AF064860.5	1.03E-35	0.19880168
use all sites	chr8	121772624	121773417	794	3	SNTB1	5.10E-34	0.29023955
use all sites	chr1	30882722	30884528	1807	3	snoU13, Y_RNA, SCARNA16, SCARNA21, U1, SCARNA17, SCARNA18, SCARNA24, SNORD112, SNORA62, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, RP4-591L5.1, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, snoU109, SNORD60	3.37E-34	0.20964563
use all sites	chr1	10825554	10826654	1101	4	snoU13, Y_RNA, SCARNA16, SCARNA21, SNORD112, SNORA62, SNORA2, U3, SNORA51, SCARNA20, SNORA67, SNORA70, SNORA77, SNORA26, U8, CASZ1, SCARNA11, SNORA31, SNORA42, SNORA40, snoU109, SNORD60	2.90E-35	0.18138101
use all sites	chr1	152131909	152133193	1285	4	snoU13, Y_RNA, SCARNA16, SNORD112, SNORA63, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, SNORA42, SNORA40, SCARNA4, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	5.45E-35	0.23156203
use all sites	chr6	12355642	12357406	1765	4	SNORA20	1.86E-32	0.19718297
use all sites	chr1	43610502	43610749	248	3	snoU13, Y_RNA, SCARNA16, U1, SCARNA17, SCARNA18, SCARNA24, SNORD112, SNORA62, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, snoU109, SNORD60, SNORD116	7.38E-34	0.20198665
use all sites	chr11	19840651	19841423	773	3	LSP1, NAV2	6.80E-34	0.20711221
use all sites	chr8	21888398	21889714	1317	4	NPM2, SNORA7	7.80E-32	0.19680303
use all sites	chr2	234118754	234119529	776	3	ATG16L1, SCARNA6, SNORD39	6.25E-34	0.18457893
use all sites	chr8	123834661	123835648	988	4	ZHX2	5.43E-34	0.19573622
use all sites	chr11	35217970	35219225	1256	4	CD44	1.81E-33	0.23477625
use all sites	chr14	100942739	100943363	625	3	WDR25	4.09E-33	0.19094421
use all sites	chr4	108814109	108815434	1326	5	SNORD50, SGMS2, SNORA11, RP11-286E11.1, snoU2_19, SNORD65	2.10E-24	0.18916995
use all sites	chr1	178845478	178846607	1130	3	snoU13, Y_RNA, SCARNA16, SNORD112, SNORA63, U3, SNORA51, SNORA25, SNORD59, SCARNA20, RALGPS2, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, snoU109, SNORD60, SNORD116	4.45E-32	0.25483433
use all sites	chr6	8343386	8344604	1219	4	SNORA20	1.87E-30	0.18524821
use all sites	chr10	98982314	98982776	463	3	SNORA17, ARHGAP19-SLIT1, ARHGAP19	4.70E-32	0.19378007
use all sites	chr4	83436999	83438155	1157	3	SNORD74, SNORA3, SNORD50, snoR442, SNORA11, TMEM150C, snoU2_19, SNORD65	9.70E-32	0.1956516
use all sites	chr3	54392776	54393507	732	3	U7, SNORD77, CACNA2D3, SNORA33, SNORA81, SNORD66, SNORD5, SNORD38, SNORD63, Metazoa_SRP, SNORA18	1.46E-32	0.21814232
use all sites	chr10	92757528	92758755	1228	4	RP11-236B18.2, SNORA17	4.80E-33	0.19380927
use all sites	chr6	128222053	128222390	338	3	SNORA8, SNORD28, THEMIS, SNORA20	1.24E-29	0.21372176
use all sites	chr9	38421899	38422794	896	3	SNORA30, IGFBPL1	3.41E-32	0.19783719
use all sites	chr15	48137483	48139022	1540	4	RP11-198M11.2	4.22E-31	0.23843156

use all sites	chr4	25871753	25872442	690	3	SMIM20, SNORD74, SNORA3, snoR442, snoU2_19, SNORD65	1.84E-28	0.21027605
use all sites	chr4	15239597	15240035	439	3	SNORA3, RP11-665G4.1	9.94E-31	0.2272664
use all sites	chr3	37615009	37615806	798	3	U7, ITGA9, SNORD77, SNORA33, SNORA81, SNORD5, SNORD38, SNORD63, Metazoa_SRP, SNORA18	5.21E-32	0.30351529
use all sites	chr2	66466834	66467093	260	3	SNORA73, SNORA64, SNORD75, SNORA12, SNORA74, SNORA19, snR65, 5S_rRNA, SNORA4, SNORD11, SNORA41, SCARNA6, SNORD39, SNORD18, SNORA36, SNORA75, ACA59, SNORA48, SNORD56, SNORA43, SNORA1, Vault	1.06E-29	0.26234382
use all sites	chr1	207112905	207114335	1431	3	snoU13, Y_RNA, SNORD112, U3, SNORA51, SNORA25, SNORA70, SNORA26, SNORA72, U8, PIGR, SNORD116	5.02E-32	0.19126275
use all sites	chr16	53078441	53079491	1051	3	SNORD111, RP11-467J12.4, SNORD33	9.15E-32	0.20698187
use all sites	chr3	9961961	9962252	292	3	IL17RC, SNORD38	6.87E-31	0.23761717
use all sites	chr3	62110823	62111806	984	3	U7, SNORD77, PTPRG, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD38, SNORD63, Metazoa_SRP, SNORA18	6.71E-31	0.2665107
use all sites	chr11	75310137	75310269	133	3	MAP6	2.52E-31	0.19386673
use all sites	chr6	116785236	116786636	1401	4	SNORA38, SNORA8, SNORD28, SNORA20	2.33E-31	0.18275341
use all sites	chr11	86437843	86438265	423	3	NA	7.64E-30	0.20608144
use all sites	chr18	21900171	21900538	368	3	SNORD23, OSBPL1A	4.57E-31	0.20458575
use all sites	chr16	56291675	56291813	139	4	GNAO1, SNORD111, SNORD33	2.46E-31	0.22834181
use all sites	chr11	58920029	58921817	1789	3	FAM111A, SNORD43	1.05E-29	0.19631217
use all sites	chr7	12728890	12729527	638	3	ARL4A	6.12E-30	-0.2237739
use all sites	chr9	3898591	3899055	465	3	GLIS3-AS1, GLIS3	1.69E-30	0.20841258
use all sites	chr12	15924004	15925024	1021	3	EPS8, snoMe28S-Am2634	1.05E-30	0.20578185
use all sites	chr5	137043431	137043540	110	3	SNORA27, SNORA68, SNORA57, 7SK, KLHL3, SNORD45, SNORD95	1.47E-30	0.21058504
use all sites	chr4	72118736	72120167	1432	3	SNORD74, SLC4A4, SNORA3, snoR442, SNORA11, SNORD42, snoU2_19, SNORD65	1.26E-30	0.20707807
use all sites	chr5	80743956	80745371	1416	4	snoZ6, SNORA27, SNORA68, RPS23P5, SNORA57, 7SK, SSBP2, SNORD45	2.90E-30	0.26443457
use all sites	chr5	37717742	37718202	461	3	WDR70, SNORA27, SNORA68, RPS23P5, SNORA57, SNORA76, 7SK, SNORD45	5.60E-30	-0.2795764
use all sites	chr6	143855543	143856248	706	3	SNORD28, SNORA20	6.52E-29	0.21952051
use all sites	chr2	70681570	70682536	967	3	SNORA73, SNORA64, SNORD75, SNORA12, SNORA74, SNORA19, snR65, 5S_rRNA, SNORA4, SNORD11, SNORA41, SCARNA6, SNORD39, SNORD18, SNORA36, SNORA75, ACA59, TGFA, SNORA48, SNORD56, SNORA43, SNORA1, Vault	1.13E-29	0.20133732
use all sites	chr5	141159440	141160944	1505	3	SNORA68, SNORA57, 7SK, SNORD45, SNORD95	7.32E-30	0.17442324
use all sites	chr5	75904722	75905788	1067	3	IQGAP2, snoZ6, SNORA27, SNORA68, RPS23P5, SNORA57, 7SK, CTD-2236F14.1, SNORD45	6.62E-17	0.22144301
use all sites	chr20	52780980	52781465	486	3	CYP24A1	3.54E-28	0.22013107
use all sites	chr1	236016383	236017325	943	3	snoU13, Y_RNA, SNORD112, SNORA25, LYST	3.24E-28	0.19765694

use all sites	chr2	166670253	166670619	367	3	SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, ACA59, SNORA48, AC009495.4, SNORA43, SNORA1, Vault	1.04E-29	0.18816915
use all sites	chr5	121804635	121805509	875	3	snoZ6, SNORA27, SNORA68, SNORA57, 7SK, CTC-210G5.1, SNORD45, SNORD95	3.01E-27	0.23748609
use all sites	chr8	60050151	60051356	1206	4	SNORA7, SNORA32	8.35E-27	0.23573717
use all sites	chr11	122179528	122180195	668	3	RP11-820L6.1	2.16E-29	0.19541611
use all sites	chr2	8453146	8453529	384	3	SNORA73, SNORA64, snR65, LINC00299, SNORA48, SNORD56, SNORA43	3.27E-29	0.19025375
use all sites	chr4	90226782	90227879	1098	4	SNORD74, SNORD50, SNORA11, GPRIN3, snoU2_19, SNORD65	3.88E-27	0.19233985
use all sites	chr22	32440788	32442147	1360	3	SLC5A1	5.76E-27	0.18651288
use all sites	chr11	46004687	46005268	582	3	PHF21A, SNORD43	2.56E-27	0.19960842
use all sites	chr3	148572849	148573450	602	3	CPB1, SNORA81, SNORD66, SNORD2, RP11-680B3.2, Metazoa_SRP, SNORA18, U4	9.06E-25	0.21510514
use all sites	chr5	124254607	124256096	1490	3	RP11-284A20.1, snoZ6, SNORA27, SNORA68, SNORA57, 7SK, SNORD45, SNORD95	3.00E-22	0.19911489
use all sites	chr14	100629365	100630425	1061	3	NA	8.27E-28	0.17203051
use all sites	chr8	101037447	101037516	70	3	SNORA32, RGS22	1.44E-28	0.23791683
use all sites	chr4	106854064	106854773	710	3	SNORD74, SNORD50, NPNT, SNORA11, snoU2_19, SNORD65	4.46E-24	0.21546397
use all sites	chr11	102576094	102577488	1395	3	MMP27	8.07E-26	0.24310727
use all sites	chr1	20504655	20505852	1198	4	snoU13, Y_RNA, SCARNA16, SCARNA21, U1, SCARNA17, SCARNA18, SNORD112, SNORA62, SNORA63, SNORA2, SNORD81, U3, SNORA51, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, ACA64, snoU109, SNORD60	1.72E-25	0.22428289
use all sites	chr8	121549480	121549740	261	3	MTBP, SNTB1	4.05E-27	0.22998515
use all sites	chr4	38942402	38943360	959	4	SNORD74, FAM114A1, SNORA3, snoR442, SNORA11, SNORD42, snoU2_19, SNORD65	1.28E-25	0.21088316
use all sites	chr11	128712280	128712932	653	4	KCNJ1	1.34E-25	0.1917719
use all sites	chr4	2363117	2363153	37	3	ZFYVE28	8.68E-28	0.20934873
use all sites	chr1	219497378	219498589	1212	3	snoU13, Y_RNA, SNORD112, U3, SNORA51, SNORA25, SNORA72, U8	2.89E-26	0.19091844
use all sites	chr2	207687061	207687959	899	4	SCARNA6, SNORD39, SNORA75, SNORA48	2.79E-25	0.22556461
use all sites	chr3	140179238	140180140	903	4	CLSTN2, SNORA81, SNORD66, SNORD2, SNORD5, RP11-68L1.1, Metazoa_SRP, SNORA18, U4	1.20E-27	0.19360227
use all sites	chr10	112515252	112515808	557	3	RBM20, SNORA17	2.06E-27	0.20131986
use all sites	chr8	123858181	123859953	1773	4	ZHX2, AC016405.2	2.12E-24	0.20848925
use all sites	chr9	37715014	37715305	292	3	SNORA30, FRMPD1, RP11-613M10.9	1.62E-26	0.19803352
use all sites	chr1	8066815	8067352	538	4	snoU13, Y_RNA, SCARNA21, SNORD112, SNORA62, U3, SNORA51, SNORA67, SNORA70, SNORA26, U8, ERRFI1, SCARNA11, SNORA31, SNORA42, SNORA40, snoU109, SNORD60	1.13E-26	0.19505773
use all sites	chr11	124938944	124940631	1688	3	SLC37A2	1.12E-24	0.21211848

use all sites	chr1	10781431	10783216	1786	3	snoU13, Y_RNA, SCARNA16, SCARNA21, SNORD112, SNORA62, SNORA2, U3, SNORA51, SCARNA20, SNORA67, SNORA70, SNORA77, SNORA26, U8, CASZ1, SCARNA11, SNORA31, SNORA42, SNORA40, snoU109, SNORD60	1.57E-25	0.18274645
use all sites	chr1	30887513	30888725	1213	3	snoU13, Y_RNA, SCARNA16, SCARNA21, U1, SCARNA17, SCARNA18, SCARNA24, RP4-591L5.2, SNORD112, SNORA62, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, snoU109, SNORD60	1.66E-22	0.1791372
use all sites	chr7	110969188	110969696	509	3	IMMP2L	1.43E-26	0.25994868
use all sites	chr3	152364824	152365794	971	4	SNORA81, SNORD66, SNORD2, Metazoa_SRP, SNORA18, U4	2.42E-23	0.2229062
use all sites	chr1	65524084	65524885	802	5	snoU13, Y_RNA, SCARNA16, U1, SCARNA18, SCARNA24, SNORD112, SNORA62, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, MIR101-1, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	5.05E-25	0.20166053
use all sites	chr2	12442966	12443780	815	3	SNORA73, AC096559.1, SNORA64, snR65, SNORD18, SNORA75, SNORA48, SNORD56, SNORA43	2.16E-26	0.30207253
use all sites	chr3	55504238	55504433	196	3	U7, SNORD77, SNORA33, SNORA81, SNORD66, SNORD5, SNORD38, SNORD63, WNT5A, Metazoa_SRP, SNORA18	1.11E-25	0.20785806
use all sites	chr2	12551974	12553424	1451	3	SNORA73, AC096559.1, SNORA64, snR65, SNORD18, SNORA75, SNORA48, SNORD56, SNORA43	1.58E-23	0.21583622
use all sites	chr2	188136244	188136528	285	3	AC007319.1, SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, SNORA48, SNORA1, Vault	5.11E-24	0.22690325
use all sites	chr9	113170902	113171134	233	4	SVEP1	6.51E-25	0.23048564
use all sites	chr1	172350469	172351357	889	3	snoU13, Y_RNA, SCARNA16, SNORD112, SNORA63, U3, SNORA51, SNORA25, SNORD59, SCARNA20, DNM3, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, PIGC, SNORD78, snoU109, SNORD60, SNORD116	7.93E-26	0.22724182
use all sites	chr3	25381567	25382277	711	3	U7, RARB, SNORA81, SNORD38, Metazoa_SRP	1.51E-25	0.20252266
use all sites	chr9	4558381	4559720	1340	3	SLC1A1, SPATA6L	3.48E-22	0.24945811
use all sites	chr1	192461180	192461723	544	3	snoU13, Y_RNA, SCARNA16, SNORD112, U3, SNORA51, SNORA25, SNORD59, SCARNA20, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, snoU109, SNORD60, SNORD116	7.65E-19	0.20910495
use all sites	chr7	2448630	2449916	1287	3	CHST12	8.58E-25	0.18440122
use all sites	chr11	125333533	125334935	1403	4	FEZ1	2.17E-25	0.19476718
use all sites	chr2	27224435	27225628	1194	3	SNORA73, MAPRE3, SNORA64, SNORA74, snR65, SCARNA6, SNORD39, SNORD18, AC013472.4, AC013472.3, SNORA75, SNORA48, SNORD56, SNORA43, SNORA1, Vault	3.46E-25	0.2357906
use all sites	chr14	100900810	100901630	821	3	WDR25	1.91E-24	0.20323027
use all sites	chr7	84122019	84122752	734	3	SEMA3A	1.08E-24	0.21517183
use all sites	chr3	99246029	99246780	752	3	U7, SNORD77, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, SNORA24, Metazoa_SRP, SNORA18	1.41E-23	0.2094015
use all sites	chr12	30952032	30953130	1099	3	LINC00941, snoMe28S-Am2634	1.05E-23	0.21262237
use all sites	chr4	138466594	138466819	226	3	SNORD65	2.40E-22	0.19531323
use all sites	chr13	42710869	42711037	169	3	SNORD36, DGKH, SNORD37	1.40E-23	0.22270488
use all sites	chr1	84872603	84874069	1467	4	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, SNORA2, DNASE2B, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	3.07E-23	0.22413141
use all sites	chr9	71155278	71156057	780	3	SNORA30, RP11-274B18.4, TMEM252	2.92E-22	0.19256575
use all sites	chr12	91576441	91578378	1938	5	SNORA9, DCN, snoMe28S-Am2634	3.74E-21	0.16892444

use all sites	chr17	12453166	12453280	115	3	SNORA69, LINC00670	1.66E-21	-0.2306214
use all sites	chr3	62200121	62200289	169	3	U7, SNORD77, PTPRG, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD38, SNORD63, Metazoa_SRP, SNORA18	1.16E-22	0.1916745
use all sites	chr4	174411279	174412482	1204	3	SNORD65	7.06E-22	0.17814612
use all sites	chr12	15315845	15316460	616	3	RERG, snoMe28S-Am2634	1.11E-20	0.19529523
use all sites	chr8	67280854	67281892	1039	3	SNORA7, SNORA32	6.41E-23	0.17945233
use all sites	chr4	47427752	47428108	357	4	SNORD74, GABRB1, SNORA3, snoR442, SNORA11, SNORD42, snoU2_19, SNORD65	4.40E-23	0.19180603
use all sites	chr8	123954739	123955611	873	3	ZHX2	4.69E-23	0.2103721
use all sites	chr1	64379958	64380706	749	3	snoU13, Y_RNA, SCARNA16, U1, SCARNA18, SCARNA24, SNORD112, SNORA62, SNORA63, SNORD46, ROR1, SNORA2, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	9.96E-22	0.21086385
use all sites	chr2	151484439	151484866	428	3	AC104777.4, 5S_rRNA, SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, ACA59, SNORA48, SNORD56, SNORA43, SNORA1, Vault	2.17E-21	0.25416434
use all sites	chr12	131526494	131527216	723	4	GPR133	1.25E-22	0.20564367
use all sites	chr6	39608096	39608711	616	3	SNORA38, SNORA8, KIF6, SNORA20	4.99E-17	0.21937287
use all sites	chr4	48108327	48108744	418	3	SNORD74, SNORA3, snoR442, SNORA11, TXK, SNORD42, snoU2_19, SNORD65	3.08E-22	0.20033834
use all sites	chr3	114105023	114106326	1304	4	ZBTB20-AS1, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, SNORD61, ZBTB20, SNORA24, Metazoa_SRP, SNORA18, U4	5.31E-21	0.20715473
use all sites	chr13	34230354	34230675	322	3	SNORD36, RP11-141M1.3, SNORD37	1.61E-20	0.19033388
use all sites	chr2	177373410	177373823	414	3	SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, ACA59, SNORA48, SNORA43, SNORA1, Vault	1.93E-19	0.18401963
use all sites	chr9	90195336	90196103	768	3	DAPK1	1.05E-19	0.19330737
use all sites	chr17	58826436	58826810	375	3	SNORA69, BCAS3	8.83E-13	0.20895397
use all sites	chr7	96133088	96133991	904	3	SHFM1	2.35E-21	0.19212112
use all sites	chr18	54622933	54624442	1510	3	WDR7	1.16E-20	0.24587942
use all sites	chr9	109426545	109426950	406	3	RP11-308N19.4	1.45E-19	0.21460959
use all sites	chr1	56038581	56038796	216	3	snoU13, Y_RNA, SCARNA16, U1, SCARNA18, SCARNA24, SNORD112, SNORA62, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, SNORA58, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, snoU109, SNORD60, SNORD116	6.66E-18	0.20004102
use all sites	chr3	108672412	108673045	634	3	U7, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, SNORD61, GUCA1C, SNORA24, Metazoa_SRP, SNORA18	9.61E-21	0.23604284
use all sites	chr1	186429967	186430584	618	3	snoU13, Y_RNA, SCARNA16, SNORD112, SNORA63, U3, SNORA51, SNORA25, SNORD59, SCARNA20, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, PDC, snoU109, SNORD60, SNORD116	2.18E-21	0.2347791
use all sites	chr3	46482949	46484379	1431	3	U7, SNORD77, SNORA33, SNORA81, SNORD5, LTF, SNORD38, SNORD63, Metazoa_SRP, SNORA18	1.68E-18	0.21398901
use all sites	chr6	116744279	116745588	1310	3	SNORA38, SNORA8, DSE, SNORD28, SNORA20	1.58E-18	0.19244886
use all sites	chr2	149793455	149793702	248	3	KIF5C, 5S_rRNA, SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, ACA59, SNORA48, SNORD56, SNORA43, SNORA1, Vault	8.88E-17	0.18993023

use all sites	chr8	127489179	127489296	118	3	RP11-103H7.1	5.54E-20	0.19193021
use all sites	chr11	16430247	16431370	1124	3	LSP1, snoMBII-202, SOX6	1.19E-18	0.21703824
use all sites	chr5	94616672	94617750	1079	3	snoZ6, SNORA27, SNORA68, SNORA57, 7SK, MCTP1, SNORD45, SNORD95	2.20E-19	0.19510877
use all sites	chr10	60274755	60275541	787	3	BICC1, SNORA71	1.40E-17	0.19854524
use all sites	chr7	70156755	70157145	391	3	AUTS2	6.83E-19	0.18538358
use all sites	chr11	130340832	130341779	948	3	ADAMTS15	2.92E-18	0.18467879
use all sites	chr10	36200347	36201305	959	3	NA	9.70E-16	0.21229466
use all sites	chr3	66533056	66534318	1263	3	U7, SNORD77, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD38, SNORD63, LRIG1, SNORA24, Metazoa_SRP, SNORA18	2.44E-17	0.21382895
use all sites	chr5	60083033	60083582	550	3	SNORA50, snoZ6, SNORA27, SNORA68, RPS23P5, SNORA57, ELOVL7, SNORA76, 7SK, SNORD45	7.32E-14	0.19806822
use all sites	chr3	8612614	8613388	775	4	LMCD1-AS1, SNORD38	1.55E-17	0.19660857
use all sites	chr2	187988552	187989001	450	3	AC007319.1, SNORA4, SNORD11, SNORD51, SNORA41, SCARNA6, SNORD39, SNORA75, SNORA48, SNORA1, Vault	2.05E-16	0.21270714
use all sites	chr6	73339270	73339537	268	3	SNORA38, SNORA8, SCARNA15, KCNQ5, SNORD28, SNORA20	3.28E-16	0.2166727
use all sites	chr4	137045738	137046811	1074	3	RP11-775H9.2, SNORD65	2.14E-12	0.17773895
use all sites	chr10	4447810	4448386	577	3	LINC00703	3.44E-15	0.17916395
use all sites	chr14	34077052	34077154	103	3	NPAS3, SNORA79	7.45E-15	0.18087706
sites at TSS2000	chr13	21275739	21277505	1767	7	IL17D, SNORA16, SNORD37	5.64E-52	0.18484952
sites at TSS2000	chr15	75039788	75041386	1599	7	CYP1A2	2.88E-53	0.20004377
sites at TSS2000	chr1	215178601	215179216	616	8	snoU13, Y_RNA, SNORD112, U3, SNORA51, SNORA25, SNORA70, KCNK2, SNORA72, U8, SNORD116	2.95E-58	0.21816757
sites at TSS2000	chr2	25551575	25552751	1177	6	SNORA73, SNORA64, SNORA74, snR65, SCARNA6, SNORD39, SNORD18, DNMT3A, MIR1301, SNORA75, SNORA48, SNORD56, SNORA43, SNORA1	7.59E-52	0.19727055
sites at TSS2000	chr13	46757395	46757637	243	4	SNORD36, LCP1, SNORD37	6.66E-55	0.21707905
sites at TSS2000	chr1	84863547	84864880	1334	7	snoU13, Y_RNA, SCARNA16, U1, SNORD112, SNORA63, SNORD46, SNORA2, DNASE2B, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	4.33E-50	0.19528599
sites at TSS2000	chr12	105322376	105323744	1369	5	SNORA9, SLC41A2	2.45E-46	0.19730917
sites at TSS2000	chr6	2750213	2752360	2148	6	MYLK4	4.29E-47	0.22118473
sites at TSS2000	chr12	79811775	79813192	1418	5	SYT1, MIR1252, SNORA9, RP1-78O14.1, snoMe28S-Am2634	1.18E-46	0.23773521
sites at TSS2000	chr12	6055201	6056642	1442	6	ANO2	1.27E-41	0.19437382
sites at TSS2000	chr6	6622864	6624448	1585	5	LY86, LY86-AS1, SNORA20	8.81E-43	0.19519653
sites at TSS2000	chr15	45802899	45803341	443	4	RP11-519G16.3, HMGN2P46, SLC30A4	2.38E-43	0.32389857

sites at TSS2000	chr11	105010284	105011637	1354	6	CARD18	4.21E-35	0.1884503
sites at TSS2000	chr12	6553092	6554051	960	5	CD27, CD27-AS1	7.78E-41	0.18968029
sites at TSS2000	chr5	146298512	146300238	1727	4	SNORA68, SNORA57, SNORD45, PPP2R2B, PPP2R2B-IT1, SNORD95	7.33E-41	0.19346385
sites at TSS2000	chr11	76744970	76745431	462	3	B3GNT6	1.26E-39	0.22185956
sites at TSS2000	chr5	60932581	60933135	555	3	SNORA50, snoZ6, SNORA27, SNORA68, RPS23P5, SNORA57, SNORA76, 7SK, SNORD45	1.05E-37	0.21717499
sites at TSS2000	chr15	39541977	39543043	1067	4	C15orf54, RP11-624L4.1	7.43E-36	0.23393044
sites at TSS2000	chr10	74652084	74653449	1366	4	OIT3, SNORA71, SNORA17	4.26E-36	0.19702601
sites at TSS2000	chr15	43513306	43513563	258	3	EPB42	1.61E-34	0.19025913
sites at TSS2000	chr6	128239611	128241185	1575	5	SNORA8, SNORD28, THEMIS, SNORA20	4.30E-33	0.21703883
sites at TSS2000	chr10	91136405	91137460	1056	4	SNORA17, LIPA	2.16E-29	0.21544411
sites at TSS2000	chr3	137851111	137852643	1533	3	SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, A4GNT, Metazoa_SRP, SNORA18, U4	6.04E-33	0.24997344
sites at TSS2000	chr4	88719777	88720640	864	6	SNORD74, SNORD50, SNORA11, snoU2_19, SNORD65	1.19E-26	0.17613958
sites at TSS2000	chr3	148581801	148583152	1352	4	CPA3, SNORA81, SNORD66, SNORD2, RP11-680B3.2, Metazoa_SRP, SNORA18, U4	3.45E-32	0.25854012
sites at TSS2000	chr2	211180001	211181383	1383	3	SCARNA6, SNORD39, SNORA75, SNORA48	5.19E-28	0.2753323
sites at TSS2000	chr17	38721675	38721942	268	3	SNORA69, CCR7	2.10E-31	0.20905431
sites at TSS2000	chr3	156840234	156841889	1656	4	SNORA81, SNORD66, SNORD2, LINC00880, SNORA18, U4	2.17E-30	0.18762665
sites at TSS2000	chr1	152131909	152133193	1285	4	snoU13, Y_RNA, SCARNA16, SNORD112, SNORA63, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, SNORA42, SNORA40, SCARNA4, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	1.09E-29	0.22397179
sites at TSS2000	chr18	21900171	21900538	368	3	SNORD23, OSBPL1A	2.29E-26	0.19778298
sites at TSS2000	chr12	58328060	58329764	1705	4	RP11-620J15.3, snoMe28S-Am2634	2.20E-20	0.19777851
sites at TSS2000	chr9	3898591	3899055	465	3	GLIS3-AS1, GLIS3	1.71E-25	0.19932396
sites at TSS2000	chr1	236016383	236017325	943	3	snoU13, Y_RNA, SNORD112, SNORA25, LYST	3.67E-24	0.19478811
sites at TSS2000	chr1	65524084	65524885	802	5	snoU13, Y_RNA, SCARNA16, U1, SCARNA18, SCARNA24, SNORD112, SNORA62, SNORA63, SNORD46, SNORA2, SNORD81, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, U8, SCARNA11, MIR101-1, SNORA31, SNORA42, SNORA40, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	3.61E-21	0.20051478
sites at TSS2000	chr9	71155278	71156057	780	3	SNORA30, RP11-274B18.4, TMEM252	1.57E-19	0.19098288
sites at TSS2000	chr4	88528262	88529234	973	3	SNORD74, SNORD50, SNORA11, RP11-742B18.1, snoU2_19, SNORD65	8.96E-19	0.22560272
sites at TSS2000	chr17	12453166	12453280	115	3	SNORA69, LINC00670	1.71E-18	-0.2272351
sites at TSS2000	chr12	91576441	91577988	1548	4	SNORA9, DCN, snoMe28S-Am2634	1.21E-18	0.18750985

sites at TSS2000	chr4	15341523	15341878	356	3	C1QTNF7, SNORA3, RP11-665G4.1	2.86E-17	0.15374642
sites at TSS2000	chr7	96133088	96133991	904	3	SHFM1	2.97E-18	0.19047381
sites at TSS2000	chr1	186429967	186430584	618	3	snoU13, Y_RNA, SCARNA16, SNORD112, SNORA63, U3, SNORA51, SNORA25, SNORD59, SCARNA20, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, PDC, snoU109, SNORD60, SNORD116	3.91E-18	0.22791304
sites at TSS2000	chr3	108672412	108673045	634	3	U7, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, SNORD61, GUCA1C, SNORA24, Metazoa_SRP, SNORA18	2.68E-16	0.22398775
sites at TSS2000	chr3	8612614	8613388	775	4	LMCD1-AS1, SNORD38	2.66E-14	0.18635531
sites at promoters	chr13	46757395	46757637	243	4	SNORD36, LCP1, SNORD37	1.53E-49	0.21240441
sites at promoters	chr12	8088741	8089207	467	5	SLC2A3, snoMe28S-Am2634	1.22E-44	0.19440806
sites at promoters	chr22	20923127	20923814	688	3	MED15	3.80E-42	-0.2156571
sites at promoters	chr17	38716802	38718229	1428	4	SNORA69, CCR7	7.50E-25	0.23133691
sites at promoters	chr6	143178039	143179316	1278	3	SNORD28, HIVEP2, SNORA20	8.01E-27	0.1777914
sites at promoters	chr9	134273620	134273881	262	3	PRRC2B	2.45E-26	0.19567164
sites at promoters	chr17	38721675	38721942	268	3	SNORA69, CCR7	3.91E-26	0.20231998
sites at promoters	chr18	21572622	21572748	127	4	TTC39C, SNORD23	1.95E-20	0.19406203
sites at enhancers	chr2	47077192	47078188	997	3	SNORA73, SNORA64, SNORD75, LINC01119, SNORA74, snR65, 5S_rRNA, SNORD11, SNORA41, SCARNA6, SNORD39, SNORD18, SNORA36, SNORA75, SNORA48, SNORD56, SNORA43, SNORA1, Vault	4.99E-51	0.19212875
sites at enhancers	chr17	3791228	3792074	847	4	CAMKK1	7.10E-50	0.22750202
sites at enhancers	chr7	4764941	4765362	422	3	FOXK1	2.16E-30	-0.2145312
sites in PMDs	chr1	215178601	215179216	616	8	snoU13, Y_RNA, SNORD112, U3, SNORA51, SNORA25, SNORA70, KCNK2, SNORA72, U8, SNORD116	5.19E-57	0.21263031
sites in PMDs	chr7	93189633	93190462	830	5	CALCR	3.02E-35	0.20671069
sites in PMDs	chr22	22383893	22385134	1242	3	PRAMENP	2.91E-35	0.21250144
sites in PMDs	chr1	152131909	152133193	1285	4	snoU13, Y_RNA, SCARNA16, SNORD112, SNORA63, U3, SNORA51, SNORA25, SNORA58, CKS1B, SCARNA20, SNORA67, U6, SNORA70, SNORA77, SNORA26, SNORA72, U8, SNORA31, SNORA42, SNORA40, SCARNA4, SNORD64, ACA64, SNORD78, snoU109, SNORD60, SNORD116	4.18E-30	0.22616413
sites in PMDs	chr5	158892858	158893397	540	4	SNORA57, SNORD45, AC008703.1, SNORD95	1.76E-20	0.18761949
sites in PMDs	chr3	99246029	99246780	752	3	U7, SNORD77, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, SNORA24, Metazoa_SRP, SNORA18	3.55E-20	0.20646385
sites in PMDs	chr3	114105023	114106326	1304	4	ZBTB20-AS1, SNORA33, SNORA81, SNORD66, SNORD2, SNORD5, SNORD63, SNORD61, ZBTB20, SNORA24, Metazoa_SRP, SNORA18, U4	7.91E-17	0.19469617
sites in PMDs	chr5	94616672	94617750	1079	3	snoZ6, SNORA27, SNORA68, SNORA57, 7SK, MCTP1, SNORD45, SNORD95	2.25E-16	0.18974943
sites in PMDs	chr6	73339270	73339537	268	3	SNORA38, SNORA8, SCARNA15, KCNQ5, SNORD28, SNORA20	5.92E-13	0.20314664
sites in PMDs	chr11	26309578	26309747	170	3	ANO3	3.09E-11	0.16518706

Table S4.5 ICRs that can be detected using EPIC array.

Chromosome	start	end	width	ICR type	Gene	ICR index	PMD	CpG island	Promoter	enhancer	Delta Beta (≤ 10 weeks' group - > 10 weeks' group)
chr1	11561497	11561710	214	mat	DISP3	ICR1	FALSE	TRUE	FALSE	FALSE	-0.0269942
chr1	11561662	11561907	246	mat	DISP3	ICR2	FALSE	TRUE	TRUE	FALSE	-0.0335496
chr1	38461540	38461879	340	pat	FHL3	ICR3	FALSE	TRUE	FALSE	FALSE	-0.030033
chr1	39249425	39249603	179	mat	RRAGC	ICR4	FALSE	TRUE	FALSE	FALSE	0.01439873
chr1	39249433	39249663	231	mat	RRAGC	ICR5	FALSE	TRUE	FALSE	FALSE	0.00684583
chr1	39249604	39254896	5293	mat	RRAGC	ICR6	FALSE	FALSE	FALSE	FALSE	0.05422418
chr1	40024971	40025410	440	mat	PPIEL	ICR7	FALSE	TRUE	TRUE	FALSE	0.03805789
chr1	40025232	40025414	183	mat	PPIEL	ICR8	FALSE	FALSE	TRUE	FALSE	0.05698946
chr1	42384310	42384389	80	pat	HIVEP3	ICR9	FALSE	TRUE	TRUE	FALSE	0.00187057
chr1	42384365	42384473	109	pat	HIVEP3	ICR10	FALSE	FALSE	FALSE	FALSE	-0.0234045
chr1	43814764	43815034	271	pat	MPL	ICR11	FALSE	TRUE	FALSE	FALSE	-0.0328412
chr1	68512539	68512776	238	mat	DIRAS3	ICR12	FALSE	FALSE	FALSE	FALSE	0.02968432
chr1	68512650	68512806	157	mat	DIRAS3	ICR13	FALSE	FALSE	FALSE	FALSE	0.02247839
chr1	68512777	68512844	68	mat	DIRAS3	ICR14	FALSE	TRUE	FALSE	FALSE	0.01467518
chr1	68512807	68512927	121	mat	DIRAS3	ICR15	FALSE	TRUE	FALSE	FALSE	0.01297799
chr1	68512845	68512970	126	mat	DIRAS3	ICR16	FALSE	FALSE	FALSE	FALSE	0.00897054
chr1	68512928	68513062	135	mat	DIRAS3	ICR17	FALSE	FALSE	FALSE	FALSE	0.02457739
chr1	68515788	68515976	189	mat	DIRAS3	ICR18	FALSE	FALSE	FALSE	FALSE	-0.0041723
chr1	68515872	68516079	208	mat	DIRAS3	ICR19	FALSE	TRUE	FALSE	FALSE	-0.0276718
chr1	68515977	68516100	124	mat	DIRAS3	ICR20	FALSE	TRUE	FALSE	FALSE	-0.0437175
chr1	68516138	68516373	236	mat	DIRAS3	ICR23	FALSE	TRUE	TRUE	FALSE	2.64E-04
chr1	68516272	68516452	181	mat	DIRAS3	ICR24	FALSE	TRUE	TRUE	FALSE	2.64E-04
chr1	68516518	68517176	659	mat	GNG12-AS1	ICR29	FALSE	FALSE	FALSE	FALSE	0.01667993
chr1	68516713	68517204	492	mat	GNG12-AS1	ICR30	FALSE	FALSE	FALSE	FALSE	0.00149469
chr1	68517177	68517254	78	mat	GNG12-AS1	ICR31	FALSE	TRUE	FALSE	FALSE	-0.0145428
chr1	68517205	68517272	68	mat	GNG12-AS1	ICR32	FALSE	TRUE	FALSE	FALSE	-0.0153951
chr1	68517255	68517655	401	mat	GNG12-AS1	ICR33	FALSE	FALSE	FALSE	FALSE	0.01277093
chr1	102308808	102312609	3802	mat	OLFM3	ICR34	TRUE	FALSE	TRUE	FALSE	-0.0131465
chr1	149147625	149148260	636	mat	LINC02591	ICR36	FALSE	TRUE	FALSE	FALSE	-0.0096063
chr1	149170064	149170351	288	mat	LINC02591	ICR38	TRUE	TRUE	FALSE	FALSE	0.01399351
chr1	152161237	152161520	284	mat	HRNR	ICR39	FALSE	FALSE	FALSE	FALSE	0.02014128
chr1	152161397	152161682	286	mat	HRNR	ICR40	FALSE	TRUE	FALSE	FALSE	-0.0263172
chr1	152161521	152161884	364	mat	HRNR	ICR41	FALSE	TRUE	TRUE	FALSE	-0.0263172
chr1	152161683	152161926	244	mat	HRNR	ICR42	FALSE	TRUE	TRUE	FALSE	-0.0018826
chr1	152161885	152162024	140	mat	HRNR	ICR43	FALSE	TRUE	FALSE	FALSE	-0.0190013
chr1	152161927	152162506	580	mat	HRNR	ICR44	FALSE	FALSE	FALSE	FALSE	-0.0278117
chr1	248100276	248100406	131	mat	OR2L13	ICR45	FALSE	TRUE	FALSE	FALSE	-0.027803
chr2	10637974	10638112	139	pat	ODC1	ICR51	TRUE	FALSE	FALSE	FALSE	-0.057213
chr2	177015044	177015124	81	mat	MIR10B	ICR52	FALSE	TRUE	FALSE	FALSE	0.00239413
chr2	177015070	177015991	922	mat	MIR10B	ICR53	FALSE	FALSE	FALSE	FALSE	0.01517312
chr2	177015125	177016036	912	mat	MIR10B	ICR54	FALSE	FALSE	FALSE	FALSE	0.01188602
chr2	177015992	177016166	175	mat	HOXD4	ICR55	TRUE	FALSE	FALSE	FALSE	0.01100587
chr2	207116070	207118252	2183	pat	ZDBF2	ICR56	FALSE	FALSE	FALSE	FALSE	-0.0119547
chr2	207116401	207118287	1887	pat	ZDBF2	ICR57	FALSE	FALSE	FALSE	FALSE	0.01826316
chr2	207118253	207127225	8973	pat	ZDBF2	ICR58	FALSE	FALSE	FALSE	FALSE	0.00504215
chr2	207118288	207127363	9076	pat	ZDBF2	ICR59	FALSE	FALSE	FALSE	FALSE	-0.0024429
chr2	207127226	207129157	1932	pat	ZDBF2	ICR60	FALSE	FALSE	FALSE	FALSE	0.0277394

chr2	207127364	207136183	8820	pat	ZDBF2	ICR61	FALSE	TRUE	FALSE	FALSE	0.05548822
chr2	241975140	241975970	831	mat	SNED1	ICR63	FALSE	FALSE	TRUE	FALSE	0.02985119
chr2	241975971	241976243	273	mat	SNED1	ICR64	FALSE	TRUE	TRUE	FALSE	0.0622198
chr3	27674461	27674613	153	mat	EOMES	ICR65	FALSE	TRUE	FALSE	FALSE	-7.94E-04
chr3	39543515	39543775	261	mat	MOBP	ICR66	FALSE	FALSE	TRUE	FALSE	-0.0196766
chr3	39543547	39543966	420	mat	MOBP	ICR67	FALSE	FALSE	TRUE	FALSE	-0.0115225
chr3	39543776	39544191	416	mat	MOBP	ICR68	FALSE	TRUE	TRUE	FALSE	-8.29E-04
chr3	42727160	42727488	329	mat	KLHL40	ICR69	FALSE	FALSE	FALSE	FALSE	0.01616523
chr3	42727425	42727842	418	mat	KLHL40	ICR70	FALSE	TRUE	FALSE	FALSE	-0.0301814
chr3	159557552	159557796	245	pat	SCHIP1	ICR71	FALSE	FALSE	TRUE	FALSE	0.12593109
chr3	159557779	159558030	252	pat	SCHIP1	ICR72	FALSE	FALSE	TRUE	FALSE	0.12287067
chr4	6107021	6107279	259	mat	JAKMIP1	ICR74	FALSE	FALSE	FALSE	FALSE	0.01130006
chr4	6107131	6107319	189	mat	JAKMIP1	ICR75	FALSE	FALSE	FALSE	FALSE	-0.0104136
chr4	6107280	6107338	59	mat	JAKMIP1	ICR76	FALSE	TRUE	FALSE	FALSE	-0.0173527
chr4	6107320	6107632	313	mat	JAKMIP1	ICR77	FALSE	TRUE	FALSE	FALSE	-0.0066835
chr4	6107339	6107648	310	mat	JAKMIP1	ICR78	FALSE	TRUE	FALSE	FALSE	-0.0106437
chr4	89618324	89618532	209	mat	HERC3	ICR79	FALSE	FALSE	TRUE	FALSE	-0.0117902
chr4	89618411	89618636	226	mat	HERC3	ICR80	FALSE	TRUE	FALSE	FALSE	0.00514378
chr4	89618533	89618666	134	mat	HERC3	ICR81	FALSE	TRUE	FALSE	FALSE	-0.0031613
chr4	89618637	89618860	224	mat	HERC3	ICR82	FALSE	TRUE	TRUE	FALSE	-0.0102708
chr4	89618667	89618981	315	mat	HERC3	ICR83	FALSE	TRUE	TRUE	FALSE	-0.0068026
chr4	89618861	89619013	153	mat	HERC3	ICR84	FALSE	TRUE	TRUE	FALSE	-0.0073434
chr4	89618982	89619022	41	mat	HERC3	ICR85	FALSE	TRUE	FALSE	FALSE	-0.0102221
chr4	89619014	89619029	16	mat	HERC3	ICR86	FALSE	TRUE	TRUE	FALSE	-0.0108097
chr4	89619023	89619037	15	mat	HERC3	ICR87	FALSE	TRUE	FALSE	FALSE	-0.0072742
chr4	89619030	89619050	21	mat	HERC3	ICR88	FALSE	FALSE	FALSE	FALSE	-0.0050212
chr4	89619038	89619052	15	mat	HERC3	ICR89	FALSE	FALSE	FALSE	FALSE	-0.0160822
chr4	89619051	89619084	34	mat	HERC3	ICR90	FALSE	FALSE	FALSE	FALSE	-0.0123039
chr4	89619053	89619235	183	mat	HERC3	ICR91	FALSE	FALSE	FALSE	FALSE	2.98E-04
chr4	155702610	155703137	528	mat	RBM46	ICR92	FALSE	TRUE	FALSE	FALSE	-0.0261184
chr4	165898707	165898834	128	pat	TRIM61	ICR93	FALSE	FALSE	TRUE	FALSE	-0.0083785
chr4	165898825	165898911	87	pat	TRIM61	ICR94	FALSE	FALSE	TRUE	FALSE	-0.0211497
chr4	165898835	165898925	91	pat	TRIM61	ICR95	FALSE	FALSE	FALSE	FALSE	-0.0022341
chr4	189376704	189395409	18706	pat	LINC01060	ICR97	TRUE	FALSE	TRUE	TRUE	-0.0171285
chr5	191600	191792	193	mat	LRRC14B	ICR98	FALSE	TRUE	TRUE	FALSE	0.03275835
chr5	191628	191805	178	mat	LRRC14B	ICR99	FALSE	TRUE	TRUE	FALSE	0.05224109
chr5	1725285	1725822	538	pat	MIR4277	ICR100	FALSE	TRUE	FALSE	FALSE	0.01965591
chr5	42991495	42991861	367	pat	LOC648987	ICR101	FALSE	FALSE	FALSE	FALSE	0.02567563
chr5	74908125	74908169	45	pat	ANKDD1B	ICR102	FALSE	FALSE	FALSE	FALSE	0.07137778
chr5	134363562	134363822	261	mat	PITX1	ICR103	FALSE	TRUE	FALSE	FALSE	0.00863398
chr5	140480597	140480871	275	mat	PCDHB3	ICR104	TRUE	FALSE	FALSE	FALSE	0.01527873
chr5	140719225	140719961	737	mat	PCDHGA1	ICR105	FALSE	FALSE	FALSE	FALSE	0.01523694
chr5	176797920	176798048	129	pat	RGS14	ICR106	FALSE	TRUE	FALSE	FALSE	-0.020172
chr6	3848904	3849189	286	mat	FAM50B	ICR107	FALSE	FALSE	FALSE	FALSE	-0.0149571
chr6	3849095	3849234	140	mat	FAM50B	ICR108	FALSE	FALSE	FALSE	FALSE	0.02417883
chr6	3849190	3849271	82	mat	FAM50B	ICR109	FALSE	TRUE	FALSE	FALSE	0.01913257
chr6	3849235	3849276	42	mat	FAM50B	ICR110	FALSE	TRUE	FALSE	FALSE	0.01731574
chr6	3849272	3849293	22	mat	FAM50B	ICR111	FALSE	TRUE	FALSE	FALSE	0.02678559
chr6	3849277	3849326	50	mat	FAM50B	ICR112	FALSE	TRUE	FALSE	FALSE	0.00974036
chr6	3849294	3849330	37	mat	FAM50B	ICR113	FALSE	TRUE	FALSE	FALSE	0.00974036

chr6	3849350	3849390	41	mat	FAM50B	ICR116	FALSE	TRUE	FALSE	FALSE	0.01439402
chr6	3849381	3849410	30	mat	FAM50B	ICR117	FALSE	TRUE	FALSE	FALSE	0.01676577
chr6	3849391	3849433	43	mat	FAM50B	ICR118	FALSE	TRUE	FALSE	FALSE	0.01913753
chr6	3849442	3849474	33	mat	FAM50B	ICR121	FALSE	TRUE	FALSE	FALSE	0.02402767
chr6	3849458	3849535	78	mat	FAM50B	ICR122	FALSE	TRUE	FALSE	FALSE	0.015723
chr6	3849475	3849541	67	mat	FAM50B	ICR123	FALSE	TRUE	FALSE	FALSE	0.00930722
chr6	3849536	3849576	41	mat	FAM50B	ICR124	FALSE	TRUE	FALSE	FALSE	0.02417215
chr6	3849542	3849687	146	mat	FAM50B	ICR125	FALSE	TRUE	TRUE	FALSE	0.03678469
chr6	3849577	3849689	113	mat	FAM50B	ICR126	FALSE	TRUE	TRUE	FALSE	0.02968018
chr6	3849688	3849701	14	mat	FAM50B	ICR127	FALSE	TRUE	FALSE	FALSE	0.02346062
chr6	3849690	3849800	111	mat	FAM50B	ICR128	FALSE	TRUE	FALSE	FALSE	0.02398205
chr6	3849801	3850105	305	mat	FAM50B	ICR130	FALSE	TRUE	TRUE	FALSE	0.04102292
chr6	17016226	17016483	258	mat	STMND1	ICR131	TRUE	TRUE	FALSE	TRUE	0.019594
chr6	28945322	28945340	19	pat	ZNF311	ICR132	FALSE	TRUE	FALSE	FALSE	-0.0602818
chr6	32116963	32117048	86	mat	PRRT1	ICR135	FALSE	TRUE	FALSE	FALSE	0.12759792
chr6	32116994	32117078	85	mat	PRRT1	ICR136	FALSE	TRUE	FALSE	FALSE	0.10399121
chr6	32117049	32117087	39	mat	PRRT1	ICR137	FALSE	TRUE	FALSE	FALSE	0.11620612
chr6	75953853	75954052	200	pat	COX7A2	ICR140	FALSE	FALSE	FALSE	FALSE	-0.0112254
chr6	110736772	110736940	169	pat	DDO	ICR141	FALSE	FALSE	FALSE	FALSE	-0.0063944
chr6	110736865	110736957	93	pat	DDO	ICR142	FALSE	FALSE	FALSE	FALSE	-0.0505826
chr6	110736941	110737052	112	pat	DDO	ICR143	FALSE	FALSE	FALSE	FALSE	-0.0233574
chr6	144328421	144328916	496	mat	HYMAI	ICR144	FALSE	FALSE	FALSE	FALSE	3.18E-04
chr6	144328482	144329051	570	mat	HYMAI	ICR145	FALSE	FALSE	FALSE	FALSE	-0.0094333
chr6	144328917	144329171	255	mat	HYMAI	ICR146	FALSE	FALSE	FALSE	FALSE	0.00848255
chr6	144329052	144329330	279	mat	HYMAI	ICR147	FALSE	TRUE	FALSE	FALSE	0.00758002
chr6	144329172	144329381	210	mat	HYMAI	ICR148	FALSE	TRUE	TRUE	FALSE	0.0048772
chr6	144329331	144329472	142	mat	HYMAI	ICR149	FALSE	TRUE	TRUE	FALSE	-4.70E-04
chr6	144329382	144329484	103	mat	HYMAI	ICR150	FALSE	TRUE	TRUE	FALSE	0.00269747
chr6	144329473	144329731	259	mat	HYMAI	ICR151	FALSE	TRUE	TRUE	FALSE	0.00730967
chr6	144329485	144329765	281	mat	HYMAI	ICR152	FALSE	TRUE	TRUE	FALSE	0.00875462
chr6	144329789	144329828	40	mat	HYMAI	ICR156	FALSE	TRUE	FALSE	FALSE	0.01532789
chr6	144329802	144329886	85	mat	PLAGL1	ICR157	FALSE	TRUE	FALSE	FALSE	0.01824924
chr6	144329829	144329908	80	mat	PLAGL1	ICR158	FALSE	TRUE	FALSE	FALSE	0.04081517
chr6	144329887	144329921	35	mat	PLAGL1	ICR159	FALSE	TRUE	FALSE	FALSE	0.06184144
chr6	144329909	144329961	53	mat	PLAGL1	ICR160	FALSE	TRUE	FALSE	FALSE	0.06539765
chr6	160023581	160023688	108	mat	SOD2	ICR161	TRUE	TRUE	FALSE	FALSE	-0.0018173
chr6	160426268	160427500	1233	mat	IGF2R	ICR162	FALSE	TRUE	FALSE	TRUE	0.02048839
chr6	170048424	170055331	6908	mat	WDR27	ICR163	FALSE	FALSE	TRUE	TRUE	-0.0010962
chr6	170055155	170057476	2322	mat	WDR27	ICR164	FALSE	FALSE	FALSE	FALSE	-0.0234872
chr7	2764129	2764245	117	pat	GNA12	ICR165	FALSE	FALSE	FALSE	FALSE	0.01206872
chr7	2802560	2802941	382	pat	GNA12	ICR166	FALSE	FALSE	FALSE	FALSE	-0.1539074
chr7	4901750	4901797	48	mat	RADIL	ICR167	FALSE	TRUE	FALSE	FALSE	0.00687742
chr7	5183992	5184154	163	mat	ZNF890P	ICR168	FALSE	TRUE	TRUE	FALSE	-0.027714
chr7	5184014	5184294	281	mat	ZNF890P	ICR169	FALSE	FALSE	TRUE	FALSE	0.00880657
chr7	16890879	16891078	200	mat	AGR3	ICR170	FALSE	TRUE	FALSE	FALSE	-0.0065208
chr7	27127448	27127575	128	mat	HOXA1	ICR171	FALSE	FALSE	FALSE	FALSE	-0.0332065
chr7	27127501	27127751	251	mat	HOXA1	ICR172	FALSE	TRUE	TRUE	FALSE	-0.0204859
chr7	27127576	27128045	470	mat	HOXA1	ICR173	FALSE	TRUE	TRUE	FALSE	0.00620227
chr7	27127752	27128168	417	mat	HOXA1	ICR174	FALSE	TRUE	FALSE	FALSE	0.00715724
chr7	27134109	27134258	150	mat	HOXA1	ICR175	FALSE	FALSE	FALSE	FALSE	-0.0710191

chr7	27134225	27134368	144	mat	HOXA1	ICR176	FALSE	TRUE	FALSE	FALSE	-0.0700157
chr7	27137922	27138395	474	mat	HOTAIRM1	ICR177	TRUE	FALSE	FALSE	FALSE	-0.0169781
chr7	27138173	27138711	539	mat	HOTAIRM1	ICR178	TRUE	FALSE	FALSE	FALSE	0.00748979
chr7	27138396	27138750	355	mat	HOTAIRM1	ICR179	TRUE	FALSE	FALSE	FALSE	-0.0075876
chr7	50849639	50849804	166	mat	GRB10	ICR180	FALSE	TRUE	FALSE	FALSE	-0.0018583
chr7	50849723	50849930	208	mat	GRB10	ICR181	FALSE	TRUE	FALSE	FALSE	0.02473193
chr7	50849805	50850563	759	mat	GRB10	ICR182	FALSE	TRUE	TRUE	FALSE	0.03684063
chr7	50849931	50850869	939	mat	GRB10	ICR183	FALSE	FALSE	TRUE	FALSE	0.01238779
chr7	50850564	50851502	939	mat	GRB10	ICR184	FALSE	FALSE	FALSE	FALSE	0.014624
chr7	94285642	94285684	43	mat	PEG10	ICR185	FALSE	TRUE	FALSE	FALSE	0.01955521
chr7	94285672	94285711	40	mat	PEG10	ICR186	FALSE	TRUE	TRUE	FALSE	0.01816816
chr7	94285685	94285744	60	mat	PEG10	ICR187	FALSE	TRUE	FALSE	FALSE	0.01362008
chr7	94285712	94285758	47	mat	PEG10	ICR188	FALSE	TRUE	FALSE	FALSE	0.00948419
chr7	94285745	94285767	23	mat	PEG10	ICR189	FALSE	TRUE	FALSE	FALSE	0.00867481
chr7	94285759	94285776	18	mat	PEG10	ICR190	FALSE	TRUE	FALSE	FALSE	0.00692326
chr7	94285777	94285813	37	mat	PEG10	ICR191	FALSE	TRUE	FALSE	FALSE	0.00275658
chr7	94285784	94285872	89	mat	PEG10	ICR192	FALSE	TRUE	FALSE	FALSE	0.00468837
chr7	94285814	94285886	73	mat	PEG10	ICR193	FALSE	TRUE	FALSE	FALSE	0.004356
chr7	94285873	94285901	29	mat	PEG10	ICR194	FALSE	TRUE	FALSE	FALSE	0.00353553
chr7	94285887	94285910	24	mat	PEG10	ICR195	FALSE	TRUE	FALSE	FALSE	0.00497923
chr7	94285902	94285941	40	mat	PEG10	ICR196	FALSE	TRUE	FALSE	FALSE	0.01364613
chr7	94285911	94285950	40	mat	PEG10	ICR197	FALSE	TRUE	FALSE	FALSE	0.01201719
chr7	94285942	94285959	18	mat	PEG10	ICR198	FALSE	TRUE	FALSE	FALSE	0.00526359
chr7	94285951	94285992	42	mat	PEG10	ICR199	FALSE	TRUE	FALSE	FALSE	0.01498443
chr7	94285960	94286085	126	mat	PEG10	ICR200	FALSE	TRUE	FALSE	FALSE	0.03012165
chr7	94285993	94286109	117	mat	PEG10	ICR201	FALSE	TRUE	FALSE	FALSE	0.02057464
chr7	94286086	94286130	45	mat	PEG10	ICR202	FALSE	TRUE	FALSE	FALSE	0.01268498
chr7	94286110	94286207	98	mat	PEG10	ICR203	FALSE	TRUE	FALSE	FALSE	0.01572465
chr7	94286131	94286218	88	mat	PEG10	ICR204	FALSE	TRUE	FALSE	FALSE	0.01740025
chr7	94286208	94286231	24	mat	PEG10	ICR205	FALSE	TRUE	FALSE	FALSE	0.01689788
chr7	94286219	94286242	24	mat	PEG10	ICR206	FALSE	TRUE	FALSE	FALSE	0.01569653
chr7	94286232	94286260	29	mat	PEG10	ICR207	FALSE	TRUE	FALSE	FALSE	0.01402347
chr7	94286243	94286262	20	mat	PEG10	ICR208	FALSE	TRUE	FALSE	FALSE	0.01235782
chr7	94286261	94286266	6	mat	PEG10	ICR209	FALSE	TRUE	FALSE	FALSE	0.01043173
chr7	94286263	94286293	31	mat	PEG10	ICR210	FALSE	TRUE	FALSE	FALSE	0.01130865
chr7	94286267	94286303	37	mat	PEG10	ICR211	FALSE	TRUE	FALSE	FALSE	0.0180029
chr7	94286294	94286342	49	mat	PEG10	ICR212	FALSE	TRUE	FALSE	FALSE	0.01947658
chr7	94286304	94286350	47	mat	PEG10	ICR213	FALSE	TRUE	FALSE	FALSE	0.01549845
chr7	94286343	94286361	19	mat	PEG10	ICR214	FALSE	TRUE	FALSE	FALSE	0.01331293
chr7	94286351	94286419	69	mat	PEG10	ICR215	FALSE	TRUE	FALSE	FALSE	0.00921407
chr7	94286362	94286430	69	mat	PEG10	ICR216	FALSE	TRUE	FALSE	FALSE	0.01264467
chr7	94286420	94286472	53	mat	PEG10	ICR217	FALSE	TRUE	FALSE	FALSE	0.01736728
chr7	94286431	94286483	53	mat	PEG10	ICR218	FALSE	TRUE	FALSE	FALSE	0.02634454
chr7	94286473	94286510	38	mat	PEG10	ICR219	FALSE	TRUE	FALSE	FALSE	0.026531
chr7	94286484	94286520	37	mat	PEG10	ICR220	FALSE	TRUE	FALSE	FALSE	0.01585093
chr7	94286511	94286525	15	mat	PEG10	ICR221	FALSE	TRUE	FALSE	FALSE	0.01556462
chr7	94286521	94286558	38	mat	PEG10	ICR222	FALSE	TRUE	FALSE	FALSE	0.01288431
chr7	94286526	94286631	106	mat	PEG10	ICR223	FALSE	TRUE	FALSE	FALSE	0.00791625
chr7	94286559	94286649	91	mat	PEG10	ICR224	FALSE	TRUE	FALSE	FALSE	0.00515613
chr7	94286632	94286668	37	mat	PEG10	ICR225	FALSE	TRUE	FALSE	FALSE	0.01863165

chr7	94286650	94286759	110	mat	PEG10	ICR226	FALSE	TRUE	FALSE	FALSE	0.01067608
chr7	94286669	94286792	124	mat	PEG10	ICR227	FALSE	TRUE	FALSE	FALSE	0.00803492
chr7	94286760	94286833	74	mat	PEG10	ICR228	FALSE	TRUE	FALSE	FALSE	0.00715038
chr7	94286793	94286935	143	mat	PEG10	ICR229	FALSE	FALSE	FALSE	FALSE	0.00158002
chr7	94286834	94286952	119	mat	PEG10	ICR230	FALSE	FALSE	FALSE	FALSE	0.01298469
chr7	94286936	94286954	19	mat	PEG10	ICR231	FALSE	FALSE	FALSE	FALSE	0.0131023
chr7	94286953	94287185	233	mat	PEG10	ICR232	FALSE	FALSE	FALSE	TRUE	6.59E-04
chr7	94286955	94287210	256	mat	PEG10	ICR233	FALSE	FALSE	FALSE	TRUE	-0.0010552
chr7	94287186	94287241	56	mat	PEG10	ICR234	FALSE	FALSE	FALSE	TRUE	-0.0069782
chr7	130130122	130130186	65	mat	MEST	ICR235	FALSE	FALSE	FALSE	FALSE	-9.02E-04
chr7	130130155	130130287	133	mat	MEST	ICR236	FALSE	FALSE	FALSE	FALSE	0.0139702
chr7	130130187	130130319	133	mat	MEST	ICR237	FALSE	FALSE	FALSE	FALSE	0.01527175
chr7	130130288	130130382	95	mat	MEST	ICR238	FALSE	TRUE	FALSE	FALSE	0.01527175
chr7	130130383	130130480	98	mat	MEST	ICR240	FALSE	TRUE	FALSE	FALSE	0.04153829
chr7	130130478	130130587	110	mat	MEST	ICR241	FALSE	TRUE	FALSE	FALSE	0.03736206
chr7	130130481	130130739	259	mat	MEST	ICR242	FALSE	TRUE	FALSE	FALSE	0.02936681
chr7	130130588	130130746	159	mat	MEST	ICR243	FALSE	TRUE	FALSE	FALSE	0.03851998
chr7	130130740	130130752	13	mat	MEST	ICR244	FALSE	TRUE	FALSE	FALSE	0.04642485
chr7	130130747	130130917	171	mat	MEST	ICR245	FALSE	TRUE	FALSE	FALSE	0.03721912
chr7	130130753	130130994	242	mat	MEST	ICR246	FALSE	TRUE	FALSE	FALSE	0.05654931
chr7	130130918	130131084	167	mat	MEST	ICR247	FALSE	TRUE	FALSE	FALSE	0.04651948
chr7	130130995	130131135	141	mat	MEST	ICR248	FALSE	TRUE	FALSE	FALSE	0.00579891
chr7	130131085	130131137	53	mat	MEST	ICR249	FALSE	TRUE	FALSE	FALSE	4.24E-04
chr7	130131136	130131145	10	mat	MEST	ICR250	FALSE	TRUE	FALSE	FALSE	0.00750833
chr7	130131138	130131188	51	mat	MEST	ICR251	FALSE	TRUE	FALSE	FALSE	0.01401562
chr7	130131146	130131257	112	mat	MEST	ICR252	FALSE	TRUE	FALSE	FALSE	0.01528655
chr7	130131189	130131267	79	mat	MEST	ICR253	FALSE	TRUE	FALSE	FALSE	-0.0024768
chr7	130131258	130131358	101	mat	MEST	ICR254	FALSE	TRUE	FALSE	FALSE	0.00302544
chr7	130131268	130131366	99	mat	MEST	ICR255	FALSE	TRUE	FALSE	FALSE	0.00608702
chr7	130131359	130131402	44	mat	MEST	ICR256	FALSE	TRUE	FALSE	FALSE	0.00362462
chr7	130131367	130131479	113	mat	MEST	ICR257	FALSE	TRUE	FALSE	FALSE	0.00880952
chr7	130131403	130131483	81	mat	MEST	ICR258	FALSE	TRUE	FALSE	FALSE	0.00470009
chr7	130131480	130131632	153	mat	MEST	ICR259	FALSE	FALSE	FALSE	FALSE	-0.0016708
chr7	130131484	130131675	192	mat	MEST	ICR260	FALSE	FALSE	FALSE	FALSE	0.00127446
chr7	130131709	130131796	88	mat	MEST	ICR264	FALSE	TRUE	FALSE	FALSE	0.01415431
chr7	130131730	130131825	96	mat	MEST	ICR265	FALSE	TRUE	FALSE	FALSE	0.02160664
chr7	130131797	130131828	32	mat	MEST	ICR266	FALSE	TRUE	FALSE	FALSE	0.02501837
chr7	130131826	130131868	43	mat	MEST	ICR267	FALSE	TRUE	FALSE	FALSE	0.01919057
chr7	130131829	130131884	56	mat	MEST	ICR268	FALSE	TRUE	FALSE	FALSE	0.01386369
chr7	130131869	130131886	18	mat	MEST	ICR269	FALSE	TRUE	FALSE	FALSE	0.01208582
chr7	130131885	130131904	20	mat	MEST	ICR270	FALSE	TRUE	FALSE	FALSE	0.01218199
chr7	130131887	130131915	29	mat	MEST	ICR271	FALSE	TRUE	FALSE	FALSE	0.0083917
chr7	130131905	130131922	18	mat	MEST	ICR272	FALSE	TRUE	FALSE	FALSE	0.00626706
chr7	130131916	130131930	15	mat	MEST	ICR273	FALSE	TRUE	TRUE	FALSE	0.0049377
chr7	130131923	130132160	238	mat	MEST	ICR274	FALSE	TRUE	TRUE	FALSE	0.0049377
chr7	130131931	130132198	268	mat	MEST	ICR275	FALSE	TRUE	TRUE	FALSE	0.00846914
chr7	130132161	130132258	98	mat	MEST	ICR276	FALSE	TRUE	FALSE	FALSE	0.0110178
chr7	130132199	130132264	66	mat	MEST	ICR277	FALSE	TRUE	FALSE	FALSE	0.01411401
chr7	130132259	130132285	27	mat	MEST	ICR278	FALSE	TRUE	FALSE	FALSE	0.01284607
chr7	130132265	130132297	33	mat	MEST	ICR279	FALSE	TRUE	FALSE	FALSE	0.00516894

chr7	130132286	130132304	19	mat	MEST	ICR280	FALSE	TRUE	FALSE	FALSE	-6.93E-04
chr7	130132298	130132318	21	mat	MEST	ICR281	FALSE	TRUE	FALSE	FALSE	0.00600222
chr7	130132305	130132359	55	mat	MEST	ICR282	FALSE	TRUE	FALSE	FALSE	0.00600222
chr7	130132319	130132418	100	mat	MEST	ICR283	FALSE	TRUE	FALSE	FALSE	0.00653012
chr7	130132360	130132421	62	mat	MEST	ICR284	FALSE	TRUE	FALSE	FALSE	0.00653012
chr7	130132419	130132452	34	mat	MEST	ICR285	FALSE	TRUE	FALSE	FALSE	0.01528182
chr7	130132422	130132789	368	mat	MEST	ICR286	FALSE	TRUE	FALSE	FALSE	0.01528182
chr7	130132790	130134487	1698	mat	MEST	ICR288	FALSE	FALSE	FALSE	FALSE	0.00122601
chr7	138348981	138349442	462	mat	SVOPL	ICR289	FALSE	TRUE	FALSE	FALSE	0.03374852
chr7	154862548	154862968	421	mat	HTR5A	ICR290	FALSE	FALSE	TRUE	FALSE	0.00342518
chr7	154862770	154863175	406	mat	HTR5A	ICR291	FALSE	TRUE	TRUE	FALSE	-0.0024016
chr7	154862969	154863243	275	mat	HTR5A	ICR292	FALSE	TRUE	TRUE	FALSE	0.00104545
chr7	154863176	154863337	162	mat	HTR5A	ICR293	FALSE	FALSE	TRUE	FALSE	0.00269746
chr7	154863244	154863380	137	mat	HTR5A	ICR294	FALSE	FALSE	TRUE	FALSE	-0.0069154
chr7	158750244	158750416	173	mat	WDR60	ICR295	FALSE	TRUE	FALSE	FALSE	-0.0665363
chr7	158750384	158750606	223	mat	WDR60	ICR296	FALSE	TRUE	FALSE	FALSE	-0.0319846
chr7	158750417	158750984	568	mat	WDR60	ICR297	FALSE	TRUE	FALSE	FALSE	0.00868713
chr7	158750607	158751183	577	mat	WDR60	ICR298	FALSE	TRUE	FALSE	FALSE	0.01325247
chr7	158750985	158751590	606	mat	WDR60	ICR299	FALSE	FALSE	FALSE	FALSE	0.01341448
chr8	1321333	1321726	394	mat	LOC286083	ICR300	FALSE	TRUE	TRUE	FALSE	-0.015285
chr8	37605552	37605782	231	mat	ERLIN2	ICR301	FALSE	FALSE	FALSE	FALSE	-0.0053056
chr8	37605717	37605935	219	mat	ERLIN2	ICR302	FALSE	FALSE	FALSE	FALSE	0.00813985
chr8	37605783	37605977	195	mat	ERLIN2	ICR303	FALSE	FALSE	FALSE	FALSE	-0.0094115
chr8	39172097	39172110	14	mat	ADAM5	ICR304	FALSE	FALSE	FALSE	FALSE	-0.0147134
chr8	39172099	39172119	21	mat	ADAM5	ICR305	FALSE	FALSE	FALSE	FALSE	-0.0141105
chr8	58055591	58056025	435	pat	LINC01606	ICR307	TRUE	FALSE	FALSE	FALSE	-0.0142405
chr8	58055876	58056112	237	pat	LINC01606	ICR308	FALSE	FALSE	FALSE	FALSE	-0.0180301
chr8	141108113	141108997	885	mat	TRAPPC9	ICR309	FALSE	TRUE	FALSE	FALSE	0.01039214
chr8	141108607	141109050	444	mat	TRAPPC9	ICR310	FALSE	TRUE	FALSE	FALSE	0.00865091
chr8	141108998	141109730	733	mat	TRAPPC9	ICR311	FALSE	TRUE	FALSE	TRUE	0.00290476
chr8	141109051	141110259	1209	mat	TRAPPC9	ICR312	FALSE	TRUE	FALSE	FALSE	-0.0018997
chr8	141109731	141110746	1016	mat	TRAPPC9	ICR313	FALSE	TRUE	TRUE	FALSE	-0.0055205
chr8	141110260	141110899	640	mat	TRAPPC9	ICR314	FALSE	TRUE	TRUE	FALSE	0.01228524
chr8	141110747	141111079	333	mat	TRAPPC9	ICR315	FALSE	FALSE	FALSE	FALSE	0.01843877
chr8	141110900	141114390	3491	mat	TRAPPC9	ICR316	FALSE	FALSE	FALSE	FALSE	0.05471638
chr9	98075481	98079128	3648	mat	FANCC	ICR317	FALSE	FALSE	FALSE	FALSE	0.01068592
chr9	98075492	98079447	3956	mat	FANCC	ICR318	FALSE	TRUE	TRUE	FALSE	0.01068592
chr9	129377854	129379318	1465	mat	LMX1B	ICR319	FALSE	FALSE	FALSE	FALSE	-0.0224214
chr9	140311437	140312445	1009	mat	EXD3	ICR320	FALSE	FALSE	FALSE	TRUE	0.01215856
chr10	27703247	27703328	82	mat	PTCHD3	ICR323	FALSE	TRUE	TRUE	FALSE	-0.0157769
chr10	27703289	27703335	47	mat	PTCHD3	ICR324	FALSE	TRUE	FALSE	FALSE	-0.0203914
chr10	27703329	27703361	33	mat	PTCHD3	ICR325	FALSE	TRUE	FALSE	FALSE	-0.0167874
chr10	27703336	27703376	41	mat	PTCHD3	ICR326	FALSE	TRUE	FALSE	FALSE	-0.011653
chr10	43844651	43846375	1725	mat	FXYD4	ICR327	FALSE	FALSE	FALSE	TRUE	0.0416833
chr10	43846376	43846573	198	mat	FXYD4	ICR328	FALSE	TRUE	FALSE	FALSE	0.06166909
chr10	43846539	43846704	166	mat	FXYD4	ICR329	FALSE	FALSE	FALSE	FALSE	0.05252212
chr10	121578384	121578638	255	mat	INPP5F	ICR330	FALSE	FALSE	FALSE	FALSE	-0.0015376
chr11	1412397	1413281	885	pat	BRSK2	ICR331	TRUE	FALSE	FALSE	FALSE	0.02007182
chr11	1413145	1413314	170	pat	BRSK2	ICR332	TRUE	FALSE	FALSE	FALSE	0.02007182
chr11	1457119	1457345	227	mat	BRSK2	ICR333	TRUE	TRUE	TRUE	FALSE	-0.0132183

chr11	2018724	2019089	366	pat	H19	ICR334	FALSE	FALSE	TRUE	FALSE	-0.015088
chr11	2019079	2019115	37	pat	H19	ICR335	FALSE	FALSE	FALSE	FALSE	-0.0231788
chr11	2019090	2019128	39	pat	H19	ICR336	FALSE	FALSE	FALSE	FALSE	-0.0018916
chr11	2019116	2019166	51	pat	H19	ICR337	FALSE	FALSE	FALSE	FALSE	0.02493984
chr11	2019167	2019451	285	pat	H19	ICR339	FALSE	FALSE	FALSE	FALSE	-0.0035511
chr11	2019436	2019567	132	pat	H19	ICR340	FALSE	FALSE	FALSE	FALSE	-0.0086394
chr11	2019452	2019586	135	pat	H19	ICR341	FALSE	FALSE	FALSE	FALSE	-0.0137277
chr11	2019606	2019625	20	pat	H19	ICR344	FALSE	FALSE	FALSE	FALSE	0.00311798
chr11	2019624	2019654	31	pat	H19	ICR345	FALSE	TRUE	FALSE	FALSE	0.0036856
chr11	2019626	2019666	41	pat	H19	ICR346	FALSE	TRUE	FALSE	FALSE	0.0036981
chr11	2019655	2019729	75	pat	H19	ICR347	FALSE	TRUE	FALSE	FALSE	0.00482198
chr11	2019667	2019731	65	pat	H19	ICR348	FALSE	TRUE	FALSE	FALSE	0.004699
chr11	2019730	2019735	6	pat	H19	ICR349	FALSE	TRUE	FALSE	FALSE	0.00362977
chr11	2019732	2019797	66	pat	H19	ICR350	FALSE	TRUE	FALSE	FALSE	0.00718415
chr11	2019736	2019822	87	pat	H19	ICR351	FALSE	TRUE	FALSE	FALSE	0.01000581
chr11	2019798	2019858	61	pat	H19	ICR352	FALSE	TRUE	FALSE	FALSE	-0.0011086
chr11	2019823	2019861	39	pat	H19	ICR353	FALSE	TRUE	FALSE	FALSE	3.77E-04
chr11	2019859	2019929	71	pat	H19	ICR354	FALSE	TRUE	FALSE	FALSE	0.01081618
chr11	2019862	2019942	81	pat	H19	ICR355	FALSE	TRUE	FALSE	FALSE	0.01639263
chr11	2019930	2020027	98	pat	H19	ICR356	FALSE	TRUE	FALSE	FALSE	0.01388795
chr11	2019943	2020029	87	pat	H19	ICR357	FALSE	TRUE	FALSE	FALSE	0.01476126
chr11	2020286	2020313	28	pat	H19	ICR358	FALSE	FALSE	FALSE	FALSE	0.0301373
chr11	2020296	2020416	121	pat	H19	ICR359	FALSE	TRUE	FALSE	FALSE	0.03130832
chr11	2020314	2020536	223	pat	H19	ICR360	FALSE	TRUE	FALSE	FALSE	0.03388416
chr11	2020417	2020548	132	pat	H19	ICR361	FALSE	TRUE	FALSE	FALSE	0.03381965
chr11	2020537	2020555	19	pat	H19	ICR362	FALSE	TRUE	FALSE	FALSE	0.00911543
chr11	2020549	2020559	11	pat	H19	ICR363	FALSE	TRUE	FALSE	FALSE	0.01779285
chr11	2020556	2021102	547	pat	H19	ICR364	FALSE	FALSE	FALSE	FALSE	0.02681653
chr11	2020560	2021242	683	pat	H19	ICR365	FALSE	FALSE	FALSE	FALSE	0.02716277
chr11	2021915	2022385	471	pat	H19	ICR369	FALSE	FALSE	FALSE	FALSE	0.03776826
chr11	2022324	2023449	1126	pat	H19	ICR370	FALSE	FALSE	TRUE	FALSE	0.03943518
chr11	2022386	2023864	1479	pat	H19	ICR371	FALSE	FALSE	TRUE	FALSE	0.02779051
chr11	2023450	2024125	676	pat	H19	ICR372	FALSE	FALSE	FALSE	FALSE	0.00283423
chr11	2023865	2025905	2041	pat	H19	ICR373	FALSE	TRUE	FALSE	FALSE	0.01231083
chr11	2153991	2154112	122	pat	INS-IGF2	ICR374	FALSE	FALSE	TRUE	FALSE	-0.0625343
chr11	2154074	2154131	58	pat	INS-IGF2	ICR375	FALSE	FALSE	TRUE	FALSE	0.00282557
chr11	2154113	2154254	142	pat	INS-IGF2	ICR376	FALSE	TRUE	TRUE	FALSE	-0.0145463
chr11	2154132	2154431	300	pat	INS-IGF2	ICR377	FALSE	TRUE	TRUE	FALSE	-0.0587348
chr11	2154255	2154637	383	pat	INS-IGF2	ICR378	FALSE	TRUE	TRUE	FALSE	-0.0696278
chr11	2154432	2154933	502	pat	INS-IGF2	ICR379	TRUE	FALSE	TRUE	FALSE	-0.0642605
chr11	2154638	2154951	314	pat	INS-IGF2	ICR380	TRUE	FALSE	TRUE	FALSE	-0.0639239
chr11	2718307	2720462	2156	mat	KCNQ1	ICR381	FALSE	FALSE	FALSE	FALSE	0.03577909
chr11	2720229	2720809	581	mat	KCNQ1	ICR382	FALSE	FALSE	FALSE	FALSE	0.0211137
chr11	2720463	2721206	744	mat	KCNQ1	ICR383	FALSE	TRUE	FALSE	FALSE	0.01270969
chr11	2720810	2721242	433	mat	KCNQ1	ICR384	FALSE	TRUE	TRUE	FALSE	0.00445009
chr11	2721207	2721247	41	mat	KCNQ1	ICR385	FALSE	TRUE	TRUE	FALSE	0.00564917
chr11	2721243	2721335	93	mat	KCNQ1	ICR386	FALSE	TRUE	FALSE	FALSE	0.00337748
chr11	2721248	2721350	103	mat	KCNQ1	ICR387	FALSE	TRUE	FALSE	FALSE	0.00646983
chr11	2721336	2721365	30	mat	KCNQ1	ICR388	FALSE	TRUE	FALSE	FALSE	0.01183387
chr11	2721351	2721382	32	mat	KCNQ1	ICR389	FALSE	TRUE	FALSE	FALSE	0.01415866

chr11	2721366	2721408	43	mat	KCNQ1	ICR390	FALSE	TRUE	FALSE	FALSE	0.01079904
chr11	2721383	2721436	54	mat	KCNQ1	ICR391	FALSE	TRUE	FALSE	FALSE	0.00743942
chr11	2721437	2721590	154	mat	KCNQ1	ICR393	FALSE	TRUE	FALSE	FALSE	0.00541369
chr11	2721480	2721609	130	mat	KCNQ1	ICR394	FALSE	TRUE	FALSE	FALSE	0.00541369
chr11	2721610	2721631	22	mat	KCNQ1	ICR396	FALSE	TRUE	FALSE	FALSE	0.00199125
chr11	2721619	2721798	180	mat	KCNQ1	ICR397	FALSE	TRUE	FALSE	FALSE	0.00229512
chr11	2721632	2721816	185	mat	KCNQ1	ICR398	FALSE	TRUE	FALSE	FALSE	0.00879943
chr11	2721799	2721856	58	mat	KCNQ1	ICR399	FALSE	TRUE	FALSE	FALSE	0.00820865
chr11	2721817	2721865	49	mat	KCNQ1	ICR400	FALSE	TRUE	FALSE	FALSE	0.00141743
chr11	2721857	2721951	95	mat	KCNQ1	ICR401	FALSE	TRUE	FALSE	FALSE	0.01365495
chr11	2721866	2722061	196	mat	KCNQ1	ICR402	FALSE	TRUE	FALSE	FALSE	0.0107415
chr11	2721952	2722072	121	mat	KCNQ1	ICR403	FALSE	TRUE	FALSE	FALSE	0.00550038
chr11	2722062	2722075	14	mat	KCNQ1	ICR404	FALSE	TRUE	FALSE	FALSE	0.00317273
chr11	2722086	2722257	172	mat	KCNQ1	ICR409	FALSE	FALSE	FALSE	FALSE	0.01280366
chr11	2722195	2722339	145	mat	KCNQ1	ICR410	FALSE	FALSE	FALSE	FALSE	0.01280366
chr11	116370272	116371329	1058	mat	BUD13	ICR411	FALSE	FALSE	FALSE	FALSE	-0.0195181
chr11	116371188	116371390	203	mat	BUD13	ICR412	FALSE	TRUE	FALSE	FALSE	-0.0258052
chr12	10095902	10096111	210	pat	CLEC12A	ICR413	FALSE	FALSE	TRUE	FALSE	0.01037707
chr12	10095997	10096117	121	pat	CLEC12A	ICR414	FALSE	TRUE	TRUE	FALSE	0.03421544
chr12	10096112	10096151	40	pat	CLEC12A	ICR415	FALSE	FALSE	FALSE	FALSE	0.04143619
chr12	10096118	10096793	676	pat	CLEC12A	ICR416	FALSE	FALSE	FALSE	FALSE	0.03643975
chr12	14926744	14926985	242	mat	H2AFJ	ICR417	FALSE	TRUE	FALSE	FALSE	0.0254312
chr12	47219737	47219796	60	mat	SLC38A4	ICR418	FALSE	TRUE	FALSE	FALSE	-0.0027257
chr12	54385436	54385525	90	pat	MIR196A2	ICR419	TRUE	FALSE	FALSE	FALSE	0.01100961
chr12	114107712	114107802	91	mat	RBM19	ICR420	TRUE	FALSE	FALSE	FALSE	0.05134331
chr12	114885621	114886360	740	mat	TBX5-AS1	ICR421	FALSE	FALSE	FALSE	FALSE	0.10421539
chr12	130821962	130822602	641	mat	PIWIL1	ICR422	FALSE	TRUE	TRUE	FALSE	-0.0131413
chr12	130822286	130822604	319	mat	PIWIL1	ICR423	FALSE	TRUE	TRUE	FALSE	-0.0494475
chr12	130824015	130824327	313	mat	PIWIL1	ICR424	FALSE	TRUE	FALSE	FALSE	-0.0012748
chr12	131488414	131488725	312	pat	ADGRD1	ICR425	FALSE	FALSE	FALSE	FALSE	0.11067218
chr12	132671062	132671472	411	mat	GALNT9	ICR426	TRUE	TRUE	FALSE	FALSE	-0.0167361
chr12	133186923	133187173	251	pat	LRCOL1	ICR428	FALSE	TRUE	TRUE	FALSE	-0.0353946
chr13	20716432	20716727	296	mat	GJA3	ICR429	FALSE	TRUE	FALSE	FALSE	0.08229996
chr13	48892551	48893173	623	mat	RB1	ICR430	FALSE	TRUE	FALSE	FALSE	0.03710241
chr13	48892948	48893375	428	mat	RB1	ICR431	FALSE	TRUE	FALSE	FALSE	0.01868113
chr13	48893174	48893726	553	mat	RB1	ICR432	FALSE	TRUE	FALSE	FALSE	0.01868113
chr13	48893376	48893967	592	mat	RB1	ICR433	FALSE	TRUE	FALSE	FALSE	0.0659916
chr13	48893727	48894213	487	mat	RB1	ICR434	FALSE	TRUE	FALSE	FALSE	0.0659916
chr13	48893968	48894381	414	mat	RB1	ICR435	FALSE	TRUE	FALSE	FALSE	0.04386351
chr13	48894214	48894594	381	mat	RB1	ICR436	FALSE	FALSE	FALSE	FALSE	0.04326377
chr13	48894382	48895240	859	mat	RB1	ICR437	FALSE	FALSE	FALSE	FALSE	0.04266404
chr13	48894595	48895317	723	mat	RB1	ICR438	FALSE	FALSE	FALSE	FALSE	0.03455672
chr13	48895241	48895477	237	mat	RB1	ICR439	FALSE	FALSE	FALSE	FALSE	0.04551055
chr13	48895318	48895969	652	mat	RB1	ICR440	FALSE	FALSE	FALSE	FALSE	0.04742687
chr13	112984602	112984839	238	mat	SPACA7	ICR441	TRUE	TRUE	FALSE	FALSE	0.05163413
chr13	112984728	112984975	248	mat	SPACA7	ICR442	FALSE	TRUE	FALSE	FALSE	0.0340349
chr13	112984840	112985462	623	mat	SPACA7	ICR443	FALSE	TRUE	FALSE	FALSE	0.02975741
chr13	112986667	112986977	311	mat	SPACA7	ICR445	TRUE	TRUE	FALSE	FALSE	0.02366208
chr13	112986927	112989199	2273	mat	SPACA7	ICR446	TRUE	FALSE	FALSE	FALSE	0.01947059
chr14	101194145	101194747	603	mat	DLK1	ICR447	TRUE	TRUE	TRUE	FALSE	0.01592086

chr14	101290195	101290716	522	pat	MEG3	ICR448	FALSE	FALSE	FALSE	FALSE	0.00613633
chr14	101290556	101290866	311	pat	MEG3	ICR449	FALSE	FALSE	FALSE	FALSE	0.02354063
chr14	101290717	101291067	351	pat	MEG3	ICR450	FALSE	TRUE	FALSE	FALSE	0.04388012
chr14	101290867	101291099	233	pat	MEG3	ICR451	FALSE	TRUE	FALSE	FALSE	0.02567808
chr14	101291068	101291135	68	pat	MEG3	ICR452	TRUE	TRUE	FALSE	FALSE	0.01108014
chr14	101291100	101291179	80	pat	MEG3	ICR453	TRUE	TRUE	FALSE	FALSE	-1.41E-04
chr14	101291136	101291287	152	pat	MEG3	ICR454	TRUE	TRUE	FALSE	FALSE	-0.0088698
chr14	101291180	101291409	230	pat	MEG3	ICR455	TRUE	TRUE	FALSE	FALSE	-0.0202626
chr14	101291288	101291439	152	pat	MEG3	ICR456	TRUE	TRUE	FALSE	FALSE	-0.0199151
chr14	101291410	101291458	49	pat	MEG3	ICR457	TRUE	TRUE	FALSE	FALSE	-0.0055996
chr14	101291440	101291468	29	pat	MEG3	ICR458	TRUE	TRUE	FALSE	FALSE	-0.0301424
chr14	101291459	101291499	41	pat	MEG3	ICR459	TRUE	TRUE	FALSE	FALSE	-0.0356417
chr14	101291469	101291570	102	pat	MEG3	ICR460	TRUE	TRUE	FALSE	FALSE	-0.0368819
chr14	101291500	101291601	102	pat	MEG3	ICR461	TRUE	TRUE	FALSE	FALSE	-0.0507084
chr14	101291571	101291686	116	pat	MEG3	ICR462	TRUE	TRUE	FALSE	FALSE	0.08248877
chr14	101291602	101291828	227	pat	MEG3	ICR463	TRUE	TRUE	FALSE	FALSE	0.05520567
chr14	101291687	101291855	169	pat	MEG3	ICR464	TRUE	TRUE	FALSE	FALSE	0.02154079
chr14	101291829	101291890	62	pat	MEG3	ICR465	TRUE	TRUE	FALSE	FALSE	0.02390233
chr14	101291856	101291932	77	pat	MEG3	ICR466	TRUE	TRUE	FALSE	FALSE	0.0323822
chr14	101291891	101291996	106	pat	MEG3	ICR467	TRUE	TRUE	FALSE	FALSE	0.04455974
chr14	101291933	101292127	195	pat	MEG3	ICR468	FALSE	TRUE	FALSE	FALSE	0.027152
chr14	101291997	101292148	152	pat	MEG3	ICR469	FALSE	TRUE	FALSE	FALSE	0.01911791
chr14	101292128	101292305	178	pat	MEG3	ICR470	FALSE	TRUE	FALSE	FALSE	0.0420261
chr14	101292149	101292391	243	pat	MEG3	ICR471	FALSE	TRUE	FALSE	FALSE	0.02877375
chr14	101292306	101292642	337	pat	MEG3	ICR472	FALSE	TRUE	TRUE	FALSE	0.03153949
chr14	101292392	101292679	288	pat	MEG3	ICR473	FALSE	TRUE	TRUE	FALSE	0.02117063
chr14	101292643	101292871	229	pat	MEG3	ICR474	TRUE	TRUE	FALSE	FALSE	0.02780027
chr14	101292680	101292966	287	pat	MEG3	ICR475	TRUE	TRUE	FALSE	FALSE	0.03652719
chr14	101292872	101293089	218	pat	MEG3	ICR476	FALSE	TRUE	FALSE	FALSE	0.02143242
chr14	101292967	101293449	483	pat	MEG3	ICR477	TRUE	FALSE	FALSE	FALSE	0.04221329
chr14	101293090	101293725	636	pat	MEG3	ICR478	TRUE	FALSE	TRUE	FALSE	0.05479394
chr14	101293450	101293855	406	pat	MEG3	ICR479	FALSE	FALSE	TRUE	FALSE	0.04839891
chr14	101293726	101294146	421	pat	MEG3	ICR480	FALSE	TRUE	FALSE	FALSE	-0.0100934
chr15	23807180	23810162	2983	mat	MIR4508	ICR481	TRUE	FALSE	TRUE	FALSE	-0.0194138
chr15	23891780	23892655	876	mat	MAGEL2	ICR482	TRUE	TRUE	FALSE	FALSE	0.02542656
chr15	23892574	23892769	196	mat	MAGEL2	ICR483	TRUE	TRUE	FALSE	FALSE	-0.0038393
chr15	23892656	23893024	369	mat	MAGEL2	ICR484	TRUE	FALSE	TRUE	FALSE	-0.0345801
chr15	23892770	23893039	270	mat	MAGEL2	ICR485	TRUE	FALSE	TRUE	FALSE	-0.0127962
chr15	23893025	23893741	717	mat	MAGEL2	ICR486	TRUE	FALSE	TRUE	FALSE	0.00639404
chr15	23893040	23894197	1158	mat	MAGEL2	ICR487	TRUE	FALSE	FALSE	FALSE	-0.002811
chr15	23931451	23932369	919	mat	NDN	ICR488	FALSE	FALSE	FALSE	FALSE	-0.0016874
chr15	23931674	23932372	699	mat	NDN	ICR489	FALSE	TRUE	FALSE	FALSE	-0.0275516
chr15	23932370	23932396	27	mat	NDN	ICR490	FALSE	TRUE	FALSE	FALSE	-0.0481362
chr15	23932373	23932411	39	mat	NDN	ICR491	FALSE	TRUE	TRUE	FALSE	-0.0370458
chr15	23932397	23932619	223	mat	NDN	ICR492	FALSE	TRUE	TRUE	FALSE	-0.0425963
chr15	23932412	23932757	346	mat	NDN	ICR493	FALSE	FALSE	TRUE	FALSE	-0.0426486
chr15	24123705	24142934	19230	mat	NDN	ICR494	TRUE	FALSE	FALSE	FALSE	-0.0573139
chr15	24142935	24347062	204128	mat	PWRN2	ICR496	FALSE	FALSE	TRUE	FALSE	0.00741519
chr15	25031239	25068753	37515	mat	SNRPN	ICR499	TRUE	FALSE	FALSE	FALSE	0.02433305
chr15	25068738	25068756	19	mat	SNRPN	ICR500	TRUE	FALSE	FALSE	FALSE	0.01062769

chr15	25068754	25068762	9	mat	SNRPN	ICR501	TRUE	FALSE	FALSE	FALSE	-0.0162141
chr15	25068757	25068768	12	mat	SNRPN	ICR502	TRUE	FALSE	FALSE	FALSE	-0.0156288
chr15	25068763	25068789	27	mat	SNRPN	ICR503	TRUE	FALSE	FALSE	FALSE	-0.0114739
chr15	25068769	25068849	81	mat	SNRPN	ICR504	TRUE	FALSE	TRUE	FALSE	-0.0216108
chr15	25068790	25069375	586	mat	SNRPN	ICR505	TRUE	FALSE	TRUE	FALSE	-0.0367209
chr15	25068850	25092068	23219	mat	SNRPN	ICR506	TRUE	FALSE	FALSE	FALSE	-0.0345191
chr15	25092069	25093365	1297	mat	SNRPN	ICR507	FALSE	FALSE	FALSE	FALSE	-0.0068927
chr15	25093244	25093455	212	mat	SNRPN	ICR508	FALSE	TRUE	FALSE	FALSE	0.01763669
chr15	25093366	25093520	155	mat	SNRPN	ICR509	FALSE	TRUE	FALSE	FALSE	0.00230859
chr15	25123287	25123490	204	mat	SNRPN	ICR510	FALSE	FALSE	FALSE	FALSE	-0.086815
chr15	25123381	25123687	307	mat	SNRPN	ICR511	FALSE	FALSE	FALSE	TRUE	-0.0158734
chr15	25123491	25123730	240	mat	SNRPN	ICR512	FALSE	FALSE	FALSE	FALSE	0.01103004
chr15	25199713	25200405	693	mat	SNRPN	ICR513	FALSE	TRUE	TRUE	FALSE	0.00150822
chr15	25200253	25200489	237	mat	SNURF	ICR514	FALSE	FALSE	FALSE	FALSE	0.00565904
chr15	25200406	25201019	614	mat	SNURF	ICR515	FALSE	TRUE	FALSE	FALSE	-0.0019246
chr15	25200490	25201223	734	mat	SNURF	ICR516	FALSE	TRUE	FALSE	FALSE	-0.0075154
chr15	25201020	25201428	409	mat	SNURF	ICR517	FALSE	FALSE	FALSE	FALSE	0.0186444
chr15	25201224	25201731	508	mat	SNURF	ICR518	TRUE	FALSE	FALSE	FALSE	0.02055113
chr15	45314789	45314932	144	mat	SORD	ICR519	FALSE	FALSE	FALSE	FALSE	-0.0124392
chr15	45314915	45314942	28	mat	SORD	ICR520	FALSE	FALSE	FALSE	FALSE	-0.0156454
chr15	45314933	45315296	364	mat	SORD	ICR521	FALSE	FALSE	FALSE	FALSE	-0.0156115
chr15	69222592	69223017	426	mat	MIR548H4	ICR522	FALSE	FALSE	TRUE	FALSE	-0.0614647
chr15	95870440	95870474	35	mat	LINC01197	ICR523	FALSE	FALSE	FALSE	FALSE	-0.0445173
chr15	99408636	99408957	322	mat	IGF1R	ICR524	FALSE	TRUE	FALSE	TRUE	-0.0228099
chr15	99408804	99409193	390	mat	IGF1R	ICR525	FALSE	FALSE	FALSE	FALSE	-0.0304041
chr15	99409194	99409410	217	mat	IGF1R	ICR527	FALSE	TRUE	FALSE	FALSE	-0.1102498
chr15	99409360	99409505	146	mat	IGF1R	ICR528	FALSE	TRUE	FALSE	TRUE	-0.1013788
chr16	3493133	3493422	290	pat	ZNF597	ICR530	FALSE	FALSE	FALSE	FALSE	-0.0147485
chr16	3493336	3493439	104	pat	ZNF597	ICR531	FALSE	TRUE	FALSE	FALSE	-0.0127018
chr16	3493423	3493533	111	pat	ZNF597	ICR532	FALSE	TRUE	TRUE	FALSE	-0.0170508
chr16	3493440	3493613	174	pat	ZNF597	ICR533	FALSE	TRUE	TRUE	FALSE	-0.019011
chr16	3493534	3493680	147	pat	NAA60	ICR534	FALSE	TRUE	FALSE	FALSE	-0.0094883
chr16	3493614	3493814	201	pat	NAA60	ICR535	FALSE	TRUE	TRUE	FALSE	-0.0066878
chr16	3493681	3493996	316	pat	NAA60	ICR536	FALSE	TRUE	FALSE	FALSE	-0.0049054
chr16	3493815	3494093	279	pat	NAA60	ICR537	FALSE	TRUE	FALSE	FALSE	-0.0102361
chr16	3493997	3494154	158	pat	NAA60	ICR538	FALSE	TRUE	FALSE	FALSE	-0.0150054
chr16	3494094	3497279	3186	pat	NAA60	ICR539	FALSE	FALSE	FALSE	FALSE	-0.0044975
chr17	6558365	6558814	450	mat	MIR4520-1	ICR541	FALSE	TRUE	FALSE	FALSE	-0.0111093
chr17	6558440	6559108	669	mat	MIR4520-1	ICR542	FALSE	TRUE	FALSE	FALSE	-0.0111093
chr17	37024020	37024168	149	mat	LASP1	ICR543	FALSE	TRUE	FALSE	FALSE	9.92E-04
chr17	37093398	37123668	30271	mat	FBXO47	ICR544	FALSE	TRUE	TRUE	FALSE	0.04774555
chr17	37123638	37123670	33	mat	FBXO47	ICR545	FALSE	TRUE	TRUE	FALSE	-0.0065006
chr17	37123669	37123710	42	mat	FBXO47	ICR546	FALSE	TRUE	FALSE	FALSE	-0.0070961
chr17	70120182	70120410	229	mat	SOX9	ICR547	FALSE	TRUE	TRUE	FALSE	0.07033218
chr17	76871734	76876039	4306	mat	TIMP2	ICR548	FALSE	FALSE	FALSE	TRUE	0.13894136
chr17	76875678	76876205	528	mat	TIMP2	ICR549	FALSE	FALSE	FALSE	FALSE	0.03139321
chr17	76876040	76876238	199	mat	TIMP2	ICR550	FALSE	TRUE	FALSE	FALSE	0.03139321
chr18	29304111	29304481	371	mat	SLC25A52	ICR551	FALSE	FALSE	TRUE	FALSE	-0.0051177
chr19	1465556	1466161	606	mat	C19orf25	ICR553	FALSE	TRUE	FALSE	FALSE	0.02886617
chr19	4784940	4785373	434	mat	FEM1A	ICR554	FALSE	FALSE	FALSE	FALSE	-0.0176829

chr19	11517079	11517259	181	pat	RGL3	ICR555	FALSE	TRUE	FALSE	FALSE	0.06724574
chr19	48999042	49000896	1855	pat	LMTK3	ICR557	FALSE	FALSE	FALSE	FALSE	-0.0393972
chr19	49001890	49002476	587	pat	LMTK3	ICR558	FALSE	TRUE	FALSE	FALSE	-0.0355784
chr19	54040774	54040817	44	mat	ZNF331	ICR559	FALSE	FALSE	FALSE	FALSE	-0.0022357
chr19	54040813	54041162	350	mat	ZNF331	ICR560	FALSE	TRUE	FALSE	FALSE	0.00674457
chr19	54040818	54041250	433	mat	ZNF331	ICR561	FALSE	TRUE	FALSE	FALSE	0.00862918
chr19	54041163	54041302	140	mat	ZNF331	ICR562	FALSE	TRUE	FALSE	FALSE	0.00153357
chr19	54041251	54041307	57	mat	ZNF331	ICR563	FALSE	TRUE	FALSE	FALSE	0.00642584
chr19	54041303	54041328	26	mat	ZNF331	ICR564	FALSE	TRUE	FALSE	FALSE	0.00912126
chr19	54041308	54041397	90	mat	ZNF331	ICR565	FALSE	TRUE	FALSE	FALSE	0.01181669
chr19	54041329	54041855	527	mat	ZNF331	ICR566	FALSE	TRUE	TRUE	FALSE	0.01037801
chr19	54041398	54041998	601	mat	ZNF331	ICR567	FALSE	TRUE	TRUE	FALSE	0.01321358
chr19	54055365	54057414	2050	mat	ZNF331	ICR568	FALSE	FALSE	FALSE	FALSE	-0.0276766
chr19	54057208	54057704	497	mat	ZNF331	ICR569	FALSE	FALSE	FALSE	FALSE	-0.0102592
chr19	54057415	54058084	670	mat	ZNF331	ICR570	FALSE	TRUE	TRUE	FALSE	0.00262951
chr19	55476665	55477652	988	mat	NLRP2	ICR571	FALSE	FALSE	FALSE	FALSE	-0.0355878
chr19	55476717	55477754	1038	mat	NLRP2	ICR572	FALSE	FALSE	FALSE	FALSE	-0.0355878
chr19	57346735	57349678	2944	mat	PEG3	ICR574	FALSE	TRUE	TRUE	FALSE	-0.0154313
chr19	57349204	57349708	505	mat	PEG3	ICR575	FALSE	FALSE	TRUE	FALSE	0.00296896
chr19	57349679	57349814	136	mat	PEG3	ICR576	FALSE	FALSE	FALSE	FALSE	0.00296896
chr19	57349709	57350003	295	mat	PEG3	ICR577	FALSE	TRUE	FALSE	FALSE	0.01742779
chr19	57349815	57350095	281	mat	PEG3	ICR578	FALSE	TRUE	FALSE	FALSE	0.01779234
chr19	57350004	57350291	288	mat	PEG3	ICR579	FALSE	TRUE	FALSE	FALSE	0.02631495
chr19	57350096	57350312	217	mat	PEG3	ICR580	FALSE	TRUE	FALSE	FALSE	0.03447302
chr19	57350846	57351321	476	mat	PEG3	ICR584	FALSE	TRUE	FALSE	FALSE	0.01034676
chr19	57351213	57351439	227	mat	PEG3	ICR585	FALSE	TRUE	FALSE	FALSE	0.01855113
chr19	57351322	57351641	320	mat	PEG3	ICR586	FALSE	TRUE	FALSE	FALSE	0.02022753
chr19	57351440	57351790	351	mat	PEG3	ICR587	FALSE	TRUE	FALSE	FALSE	0.01409837
chr19	57351642	57352013	372	mat	PEG3	ICR588	FALSE	TRUE	FALSE	FALSE	0.01388108
chr19	57351791	57352017	227	mat	PEG3	ICR589	FALSE	TRUE	FALSE	FALSE	0.01326499
chr19	57352021	57352133	113	mat	ZIM2	ICR592	FALSE	TRUE	TRUE	FALSE	0.01192555
chr19	57352074	57352141	68	mat	ZIM2	ICR593	FALSE	TRUE	TRUE	FALSE	0.01519245
chr19	57352134	57352175	42	mat	ZIM2	ICR594	FALSE	TRUE	FALSE	FALSE	0.02191462
chr19	57352142	57352184	43	mat	ZIM2	ICR595	FALSE	TRUE	FALSE	FALSE	0.02014945
chr19	57352176	57352251	76	mat	MIMT1	ICR596	FALSE	TRUE	FALSE	FALSE	0.01249448
chr19	57352185	57352268	84	mat	MIMT1	ICR597	FALSE	TRUE	FALSE	FALSE	0.01374629
chr19	57352252	57352464	213	mat	MIMT1	ICR598	FALSE	FALSE	FALSE	FALSE	0.01384518
chr19	57352269	57352518	250	mat	MIMT1	ICR599	FALSE	FALSE	FALSE	FALSE	0.02561124
chr19	57352465	57352541	77	mat	MIMT1	ICR600	FALSE	FALSE	FALSE	FALSE	0.02382017
chr19	57352519	57352583	65	mat	MIMT1	ICR601	FALSE	FALSE	FALSE	FALSE	0.00546083
chr19	57352542	57352656	115	mat	MIMT1	ICR602	FALSE	FALSE	FALSE	FALSE	0.02003304
chr19	57352584	57352685	102	mat	MIMT1	ICR603	FALSE	FALSE	FALSE	FALSE	0.03709294
chr19	57352657	57352728	72	mat	MIMT1	ICR604	FALSE	FALSE	FALSE	FALSE	0.04647395
chr19	57352686	57352745	60	mat	MIMT1	ICR605	FALSE	FALSE	FALSE	FALSE	0.05302923
chr19	57352729	57352784	56	mat	MIMT1	ICR606	FALSE	FALSE	FALSE	FALSE	0.03938404
chr19	57352746	57352806	61	mat	MIMT1	ICR607	FALSE	FALSE	FALSE	FALSE	0.02175228
chr19	57352785	57353127	343	mat	MIMT1	ICR608	FALSE	FALSE	FALSE	FALSE	0.00607622
chr19	58861502	58862062	561	mat	A1BG-AS1	ICR609	FALSE	TRUE	FALSE	FALSE	0.03652569
chr20	30134929	30135107	179	mat	HM13	ICR610	FALSE	FALSE	FALSE	FALSE	0.01201711
chr20	30134973	30135123	151	mat	HM13	ICR611	FALSE	FALSE	FALSE	FALSE	0.0187073

chr20	30135108	30135148	41	mat	HM13	ICR612	FALSE	TRUE	FALSE	FALSE	0.01418526
chr20	30135124	30135157	34	mat	HM13	ICR613	FALSE	TRUE	FALSE	FALSE	0.00885912
chr20	30135149	30135176	28	mat	HM13	ICR614	FALSE	TRUE	FALSE	FALSE	0.0036599
chr20	30135158	30135290	133	mat	HM13	ICR615	FALSE	TRUE	TRUE	FALSE	-6.49E-04
chr20	30135177	30135361	185	mat	HM13	ICR616	FALSE	FALSE	TRUE	FALSE	-0.0050848
chr20	30135291	30138142	2852	mat	HM13	ICR617	FALSE	FALSE	FALSE	FALSE	0.01786353
chr20	36148457	36148614	158	mat	BLCAP	ICR618	FALSE	FALSE	FALSE	FALSE	0.01420337
chr20	36148604	36148619	16	mat	BLCAP	ICR619	FALSE	FALSE	FALSE	FALSE	0.01959224
chr20	36148615	36148641	27	mat	BLCAP	ICR620	FALSE	FALSE	FALSE	FALSE	0.02328239
chr20	36148620	36148671	52	mat	BLCAP	ICR621	FALSE	FALSE	FALSE	FALSE	0.02469328
chr20	36148642	36148678	37	mat	BLCAP	ICR622	FALSE	TRUE	FALSE	FALSE	0.03513404
chr20	36148672	36148698	27	mat	BLCAP	ICR623	FALSE	TRUE	FALSE	FALSE	0.03986153
chr20	36148679	36148737	59	mat	BLCAP	ICR624	FALSE	TRUE	FALSE	FALSE	0.03771416
chr20	36148699	36148766	68	mat	BLCAP	ICR625	FALSE	TRUE	FALSE	FALSE	0.03051365
chr20	36148738	36148774	37	mat	BLCAP	ICR626	FALSE	TRUE	FALSE	FALSE	0.01904994
chr20	36148767	36148778	12	mat	BLCAP	ICR627	FALSE	TRUE	FALSE	FALSE	0.01647276
chr20	36148775	36148790	16	mat	BLCAP	ICR628	FALSE	TRUE	FALSE	FALSE	0.02015395
chr20	36148779	36148802	24	mat	BLCAP	ICR629	FALSE	TRUE	FALSE	FALSE	0.01762067
chr20	36148803	36148927	125	mat	BLCAP	ICR630	FALSE	TRUE	FALSE	FALSE	0.01185574
chr20	36148860	36148953	94	mat	BLCAP	ICR631	FALSE	TRUE	FALSE	FALSE	0.01696613
chr20	36148928	36148960	33	mat	BLCAP	ICR632	FALSE	TRUE	FALSE	FALSE	0.02559503
chr20	36148954	36148993	40	mat	BLCAP	ICR633	FALSE	TRUE	FALSE	FALSE	0.02519966
chr20	36148961	36149012	52	mat	BLCAP	ICR634	FALSE	TRUE	FALSE	FALSE	0.03521584
chr20	36148994	36149021	28	mat	BLCAP	ICR635	FALSE	TRUE	FALSE	FALSE	0.03521584
chr20	36149053	36149111	59	mat	BLCAP	ICR638	FALSE	TRUE	FALSE	FALSE	0.02284227
chr20	36149081	36149118	38	mat	BLCAP	ICR639	FALSE	TRUE	FALSE	FALSE	0.03008924
chr20	36149112	36149120	9	mat	BLCAP	ICR640	FALSE	TRUE	FALSE	FALSE	0.02739178
chr20	36149119	36149184	66	mat	BLCAP	ICR641	FALSE	TRUE	FALSE	FALSE	0.01635326
chr20	36149121	36149187	67	mat	BLCAP	ICR642	FALSE	TRUE	FALSE	FALSE	0.02838872
chr20	36149185	36149193	9	mat	BLCAP	ICR643	FALSE	TRUE	FALSE	FALSE	0.04044761
chr20	36149188	36149230	43	mat	BLCAP	ICR644	FALSE	TRUE	FALSE	FALSE	0.03435242
chr20	36149194	36149451	258	mat	BLCAP	ICR645	FALSE	TRUE	FALSE	FALSE	0.03755935
chr20	36149231	36149454	224	mat	BLCAP	ICR646	FALSE	TRUE	FALSE	FALSE	0.0268743
chr20	36149452	36149655	204	mat	BLCAP	ICR647	FALSE	TRUE	FALSE	FALSE	0.01238705
chr20	36149455	36149705	251	mat	BLCAP	ICR648	FALSE	TRUE	TRUE	FALSE	0.04677957
chr20	36149656	36149749	94	mat	NNAT	ICR649	FALSE	TRUE	TRUE	FALSE	0.04935779
chr20	36149706	36150060	355	mat	NNAT	ICR650	FALSE	TRUE	TRUE	FALSE	0.04619956
chr20	36149750	36151002	1253	mat	NNAT	ICR651	FALSE	FALSE	FALSE	FALSE	0.04187495
chr20	42142236	42142450	215	mat	L3MBTL1	ICR652	FALSE	FALSE	FALSE	FALSE	0.00930712
chr20	42142417	42142483	67	mat	L3MBTL1	ICR653	FALSE	FALSE	FALSE	FALSE	0.00907771
chr20	42142451	42142493	43	mat	L3MBTL1	ICR654	FALSE	FALSE	FALSE	FALSE	0.00353741
chr20	42142484	42142558	75	mat	L3MBTL1	ICR655	FALSE	FALSE	FALSE	FALSE	0.00378908
chr20	42142494	42142595	102	mat	L3MBTL1	ICR656	FALSE	FALSE	FALSE	FALSE	0.00274637
chr20	42142559	42142670	112	mat	L3MBTL1	ICR657	FALSE	FALSE	FALSE	FALSE	-4.25E-04
chr20	42142596	42142672	77	mat	L3MBTL1	ICR658	FALSE	FALSE	FALSE	FALSE	-0.0133838
chr20	42142671	42142698	28	mat	L3MBTL1	ICR659	FALSE	TRUE	FALSE	FALSE	-0.0237347
chr20	42142673	42142750	78	mat	L3MBTL1	ICR660	FALSE	TRUE	FALSE	FALSE	-0.0164913
chr20	42142699	42142765	67	mat	L3MBTL1	ICR661	FALSE	FALSE	FALSE	FALSE	-0.0027208
chr20	42142751	42142783	33	mat	L3MBTL1	ICR662	FALSE	FALSE	FALSE	FALSE	7.10E-04
chr20	42142766	42142846	81	mat	L3MBTL1	ICR663	FALSE	FALSE	FALSE	FALSE	-6.98E-04

chr20	42142784	42142851	68	mat	L3MBTL1	ICR664	FALSE	FALSE	FALSE	FALSE	0.00172173
chr20	42142852	42142946	95	mat	L3MBTL1	ICR666	FALSE	FALSE	FALSE	FALSE	0.02741077
chr20	42142897	42142994	98	mat	L3MBTL1	ICR667	FALSE	FALSE	FALSE	FALSE	0.02741077
chr20	42142947	42143014	68	mat	L3MBTL1	ICR668	FALSE	FALSE	FALSE	FALSE	-0.0321437
chr20	42142995	42143044	50	mat	L3MBTL1	ICR669	FALSE	FALSE	FALSE	FALSE	-0.0394682
chr20	42143015	42143095	81	mat	L3MBTL1	ICR670	FALSE	FALSE	FALSE	FALSE	0.00240177
chr20	42143045	42143173	129	mat	L3MBTL1	ICR671	FALSE	FALSE	TRUE	FALSE	0.02750593
chr20	42143096	42143210	115	mat	L3MBTL1	ICR672	FALSE	FALSE	TRUE	FALSE	0.02605875
chr20	42143174	42143398	225	mat	L3MBTL1	ICR673	FALSE	FALSE	FALSE	FALSE	0.01473589
chr20	42143211	42143488	278	mat	L3MBTL1	ICR674	FALSE	FALSE	FALSE	FALSE	0.00762548
chr20	42143399	42143501	103	mat	L3MBTL1	ICR675	FALSE	TRUE	FALSE	FALSE	0.00937682
chr20	44782039	44784897	2859	mat	CDH22	ICR676	FALSE	FALSE	FALSE	FALSE	0.01549235
chr20	44838776	44839186	411	mat	CDH22	ICR677	FALSE	FALSE	FALSE	FALSE	-0.0016874
chr20	57414039	57414161	123	pat	GNAS-AS1	ICR678	FALSE	FALSE	FALSE	FALSE	-0.0019011
chr20	57414059	57414216	158	pat	GNAS-AS1	ICR679	FALSE	FALSE	FALSE	FALSE	0.0193447
chr20	57414162	57414273	112	pat	GNAS-AS1	ICR680	FALSE	FALSE	FALSE	FALSE	0.05069935
chr20	57414274	57414406	133	pat	GNAS-AS1	ICR681	FALSE	FALSE	FALSE	FALSE	0.03323491
chr20	57414351	57414528	178	pat	GNAS-AS1	ICR682	FALSE	FALSE	FALSE	FALSE	0.03323491
chr20	57414407	57414538	132	pat	GNAS-AS1	ICR683	FALSE	FALSE	FALSE	FALSE	0.05947765
chr20	57414529	57414577	49	pat	GNAS-AS1	ICR684	FALSE	FALSE	FALSE	FALSE	0.04793272
chr20	57414539	57414595	57	pat	GNAS-AS1	ICR685	FALSE	FALSE	FALSE	FALSE	0.03327886
chr20	57414578	57414883	306	pat	GNAS-AS1	ICR686	FALSE	FALSE	TRUE	FALSE	0.0316727
chr20	57414596	57414959	364	pat	GNAS-AS1	ICR687	FALSE	FALSE	TRUE	FALSE	0.03242408
chr20	57414884	57415143	260	pat	GNAS	ICR688	FALSE	TRUE	FALSE	FALSE	0.03984675
chr20	57414960	57415176	217	pat	GNAS	ICR689	FALSE	TRUE	FALSE	FALSE	0.03984675
chr20	57415144	57415376	233	pat	GNAS	ICR690	FALSE	TRUE	FALSE	FALSE	0.02979022
chr20	57415177	57415696	520	pat	GNAS	ICR691	FALSE	TRUE	FALSE	FALSE	0.0488785
chr20	57415377	57416220	844	pat	GNAS	ICR692	FALSE	TRUE	TRUE	FALSE	0.05322346
chr20	57415697	57416505	809	pat	GNAS	ICR693	FALSE	TRUE	TRUE	FALSE	0.0178953
chr20	57416221	57416887	667	pat	GNAS	ICR694	FALSE	TRUE	TRUE	FALSE	-0.0109125
chr20	57416506	57417151	646	pat	GNAS	ICR695	FALSE	TRUE	TRUE	FALSE	0.01155404
chr20	57416888	57417232	345	pat	GNAS	ICR696	FALSE	TRUE	FALSE	FALSE	0.00691536
chr20	57417152	57418014	863	pat	GNAS	ICR697	FALSE	FALSE	FALSE	FALSE	0.02402853
chr20	57417233	57420941	3709	pat	GNAS	ICR698	FALSE	FALSE	FALSE	FALSE	-0.0020079
chr20	57425515	57425985	471	mat	GNAS	ICR699	FALSE	FALSE	TRUE	FALSE	-0.0052704
chr20	57425994	57426137	144	mat	GNAS	ICR702	FALSE	TRUE	FALSE	FALSE	0.05579752
chr20	57426131	57426214	84	mat	GNAS	ICR703	FALSE	TRUE	FALSE	FALSE	0.04822903
chr20	57426138	57426263	126	mat	GNAS	ICR704	FALSE	TRUE	FALSE	FALSE	0.04066054
chr20	57426274	57426367	94	mat	GNAS	ICR707	FALSE	TRUE	FALSE	FALSE	0.04491596
chr20	57426322	57426373	52	mat	GNAS	ICR708	FALSE	TRUE	FALSE	FALSE	-2.71E-04
chr20	57426368	57426382	15	mat	GNAS	ICR709	FALSE	TRUE	FALSE	FALSE	-0.0434822
chr20	57426374	57426390	17	mat	GNAS	ICR710	FALSE	TRUE	FALSE	FALSE	-0.0403319
chr20	57426383	57426394	12	mat	GNAS	ICR711	FALSE	TRUE	FALSE	FALSE	-0.0343356
chr20	57426391	57426419	29	mat	GNAS	ICR712	FALSE	TRUE	FALSE	FALSE	-0.0252573
chr20	57426395	57426424	30	mat	GNAS	ICR713	FALSE	TRUE	FALSE	FALSE	-0.0210011
chr20	57426425	57426544	120	mat	GNAS	ICR715	FALSE	TRUE	FALSE	FALSE	-0.0165731
chr20	57426538	57426569	32	mat	GNAS	ICR716	FALSE	TRUE	FALSE	FALSE	-0.0165731
chr20	57426759	57426800	42	mat	GNAS	ICR723	FALSE	TRUE	FALSE	FALSE	-0.0371853
chr20	57426789	57426834	46	mat	GNAS	ICR724	FALSE	TRUE	FALSE	FALSE	-0.0394562
chr20	57426801	57426857	57	mat	GNAS	ICR725	FALSE	TRUE	FALSE	FALSE	-0.0390011

chr20	57426835	57426930	96	mat	GNAS	ICR726	FALSE	TRUE	FALSE	FALSE	-0.0286434
chr20	57426858	57426949	92	mat	GNAS	ICR727	FALSE	TRUE	FALSE	FALSE	-0.0207016
chr20	57426931	57426978	48	mat	GNAS	ICR728	FALSE	TRUE	FALSE	FALSE	-0.0293724
chr20	57426950	57427009	60	mat	GNAS	ICR729	FALSE	TRUE	FALSE	FALSE	-0.0390112
chr20	57426979	57427016	38	mat	GNAS	ICR730	FALSE	TRUE	FALSE	FALSE	-0.0365857
chr20	57427010	57427029	20	mat	GNAS	ICR731	FALSE	TRUE	FALSE	FALSE	-0.033473
chr20	57427017	57427045	29	mat	GNAS	ICR732	FALSE	TRUE	FALSE	FALSE	-0.0415941
chr20	57427030	57427102	73	mat	GNAS	ICR733	FALSE	TRUE	FALSE	FALSE	-0.0497442
chr20	57427146	57427172	27	mat	GNAS	ICR737	FALSE	TRUE	FALSE	FALSE	-0.0198294
chr20	57427170	57427209	40	mat	GNAS	ICR738	FALSE	TRUE	FALSE	FALSE	-0.0215628
chr20	57427173	57427236	64	mat	GNAS	ICR739	FALSE	TRUE	FALSE	FALSE	-0.0264617
chr20	57427210	57427273	64	mat	GNAS	ICR740	FALSE	TRUE	FALSE	FALSE	-0.0296274
chr20	57427237	57427411	175	mat	GNAS	ICR741	FALSE	TRUE	FALSE	FALSE	-0.0196224
chr20	57427274	57427425	152	mat	GNAS	ICR742	FALSE	TRUE	FALSE	FALSE	-0.0148909
chr20	57427412	57427442	31	mat	GNAS	ICR743	FALSE	TRUE	FALSE	FALSE	0.01051162
chr20	57427426	57427471	46	mat	GNAS	ICR744	FALSE	TRUE	FALSE	FALSE	0.00851369
chr20	57427443	57427482	40	mat	GNAS	ICR745	FALSE	TRUE	FALSE	FALSE	-0.0124337
chr20	57427472	57427492	21	mat	GNAS	ICR746	FALSE	TRUE	FALSE	FALSE	-0.0144133
chr20	57427483	57427494	12	mat	GNAS	ICR747	FALSE	TRUE	FALSE	FALSE	-0.0219424
chr20	57427493	57427502	10	mat	GNAS	ICR748	FALSE	TRUE	FALSE	FALSE	-0.0229146
chr20	57427495	57427555	61	mat	GNAS	ICR749	FALSE	TRUE	FALSE	FALSE	-0.0254056
chr20	57427503	57427641	139	mat	GNAS	ICR750	FALSE	TRUE	FALSE	FALSE	-0.0283277
chr20	57427556	57427649	94	mat	GNAS	ICR751	FALSE	TRUE	FALSE	FALSE	0.01477104
chr20	57427642	57427729	88	mat	GNAS	ICR752	FALSE	TRUE	FALSE	FALSE	0.0249909
chr20	57427650	57427737	88	mat	GNAS	ICR753	FALSE	TRUE	FALSE	FALSE	0.02092105
chr20	57427730	57427761	32	mat	GNAS	ICR754	FALSE	TRUE	FALSE	FALSE	0.04057611
chr20	57427738	57427820	83	mat	GNAS	ICR755	FALSE	TRUE	TRUE	FALSE	0.03812576
chr20	57427762	57427829	68	mat	GNAS	ICR756	FALSE	TRUE	TRUE	FALSE	0.04380943
chr20	57427821	57427941	121	mat	GNAS	ICR757	FALSE	TRUE	FALSE	FALSE	0.0589746
chr20	57427830	57427972	143	mat	GNAS	ICR758	FALSE	TRUE	FALSE	FALSE	0.07273493
chr20	57427942	57428031	90	mat	GNAS	ICR759	FALSE	FALSE	FALSE	FALSE	0.0572523
chr20	57429277	57430312	1036	mat	GNAS	ICR760	FALSE	TRUE	TRUE	FALSE	0.013693
chr20	57429858	57430662	805	mat	GNAS	ICR761	FALSE	TRUE	TRUE	FALSE	0.01090011
chr20	57430313	57431201	889	mat	GNAS	ICR762	FALSE	TRUE	FALSE	FALSE	0.01275199
chr20	57430663	57431302	640	mat	GNAS	ICR763	FALSE	TRUE	FALSE	FALSE	0.02018964
chr20	57463270	57463329	60	mat	GNAS	ICR764	FALSE	FALSE	FALSE	FALSE	0.06725492
chr20	57463325	57463354	30	mat	GNAS	ICR765	FALSE	FALSE	FALSE	FALSE	0.07205776
chr20	57463330	57463356	27	mat	GNAS	ICR766	FALSE	FALSE	FALSE	FALSE	0.08565816
chr20	57463355	57463396	42	mat	GNAS	ICR767	FALSE	FALSE	FALSE	FALSE	0.09852517
chr20	57463357	57463454	98	mat	GNAS	ICR768	FALSE	FALSE	FALSE	FALSE	0.1182955
chr20	57463397	57463502	106	mat	GNAS	ICR769	FALSE	FALSE	FALSE	FALSE	0.12485216
chr20	57463503	57463529	27	mat	GNAS	ICR771	FALSE	TRUE	FALSE	FALSE	0.12367546
chr20	57463527	57463571	45	mat	GNAS	ICR772	FALSE	TRUE	FALSE	FALSE	0.12012512
chr20	57463530	57463614	85	mat	GNAS	ICR773	FALSE	TRUE	FALSE	FALSE	0.11249574
chr20	57463572	57463652	81	mat	GNAS	ICR774	FALSE	TRUE	FALSE	FALSE	0.11759547
chr20	57463615	57463724	110	mat	GNAS	ICR775	FALSE	TRUE	FALSE	FALSE	0.14647088
chr20	57463653	57463762	110	mat	GNAS	ICR776	FALSE	TRUE	FALSE	FALSE	0.14880907
chr20	57463725	57463766	42	mat	GNAS	ICR777	FALSE	TRUE	FALSE	FALSE	0.10465852
chr20	57463763	57463774	12	mat	GNAS	ICR778	FALSE	TRUE	FALSE	FALSE	0.0708519
chr20	57463767	57463782	16	mat	GNAS	ICR779	FALSE	TRUE	FALSE	FALSE	0.0614666

chr20	57463775	57463786	12	mat	GNAS	ICR780	FALSE	TRUE	FALSE	FALSE	0.08602877
chr20	57463783	57463899	117	mat	GNAS	ICR781	FALSE	TRUE	TRUE	FALSE	0.11401714
chr20	57463787	57463902	116	mat	GNAS	ICR782	FALSE	TRUE	TRUE	FALSE	0.11347561
chr20	57463900	57463905	6	mat	GNAS	ICR783	FALSE	TRUE	TRUE	FALSE	0.10950788
chr20	57463903	57463924	22	mat	GNAS	ICR784	FALSE	TRUE	FALSE	FALSE	0.11820392
chr20	57463906	57463983	78	mat	GNAS	ICR785	FALSE	TRUE	FALSE	FALSE	0.11001229
chr20	57463925	57463990	66	mat	GNAS	ICR786	FALSE	TRUE	FALSE	FALSE	0.11304997
chr20	57463984	57463999	16	mat	GNAS	ICR787	FALSE	TRUE	FALSE	FALSE	0.12731695
chr20	57464571	57464969	399	mat	GNAS	ICR792	FALSE	FALSE	FALSE	FALSE	0.12244321
chr20	57464742	57464972	231	mat	GNAS	ICR793	FALSE	FALSE	FALSE	FALSE	0.10911334
chr20	57464970	57465122	153	mat	GNAS	ICR794	FALSE	TRUE	FALSE	FALSE	0.08972053
chr20	57464973	57465124	152	mat	GNAS	ICR795	FALSE	TRUE	FALSE	FALSE	0.08641373
chr20	57465123	57465131	9	mat	GNAS	ICR796	FALSE	TRUE	FALSE	FALSE	0.08595327
chr20	57465125	57465138	14	mat	GNAS	ICR797	FALSE	FALSE	FALSE	FALSE	0.09396319
chr20	57465132	57465174	43	mat	GNAS	ICR798	FALSE	FALSE	FALSE	FALSE	0.11238569
chr20	57465139	57465438	300	mat	GNAS	ICR799	FALSE	FALSE	FALSE	FALSE	0.12593226
chr20	57465175	57465444	270	mat	GNAS	ICR800	FALSE	FALSE	FALSE	FALSE	0.12170898
chr20	62570072	62570695	624	mat	UCKL1	ICR801	FALSE	TRUE	FALSE	FALSE	0.0673257
chr21	40757691	40757898	208	mat	WRB	ICR802	FALSE	TRUE	FALSE	FALSE	0.00243194
chr21	40757750	40758207	458	mat	WRB	ICR803	FALSE	FALSE	FALSE	FALSE	0.01438679
chr21	47715006	47716528	1523	mat	YBEY	ICR804	FALSE	FALSE	FALSE	FALSE	-0.0736649
chr21	47716443	47717405	963	mat	YBEY	ICR805	FALSE	FALSE	FALSE	FALSE	-0.1170913
chr22	31318240	31318372	133	mat	MORC2-AS1	ICR806	FALSE	FALSE	TRUE	FALSE	-0.0369942
chr22	42077939	42078329	391	mat	SNU13	ICR807	FALSE	FALSE	FALSE	FALSE	-0.0114743
chr22	42078365	42078566	202	mat	SNU13	ICR808	FALSE	FALSE	TRUE	FALSE	-0.0081365
chr22	42078388	42078706	319	mat	SNU13	ICR809	FALSE	FALSE	TRUE	FALSE	-0.0022048
chr22	42078567	42078722	156	mat	SNU13	ICR810	FALSE	FALSE	TRUE	FALSE	0.00412059
chr22	42548356	42548791	436	pat	TCF20	ICR811	FALSE	TRUE	FALSE	FALSE	-0.0384239
chr22	42548783	42548867	85	pat	TCF20	ICR812	FALSE	TRUE	FALSE	FALSE	-0.0360332

Table S4.6 Cohort clinical information

Variables	≤ 10 weeks' (n=48)	≥ 10 weeks' (n=83)	P values
Fetal sex (F:M)	27:21	36:47	ns (0.204)
Gestational age (weeks)	7.83 (6-10)	15.98 (11-23)	< 2.2e-16
Trimester (T1:T2)	48:0	15:68	< 2.2e-16

Table S4.7 DMR overlapped genes (Minimum probe number for each DMR is 3).

DMRs	Overlapped genes	Overlapped promoter of genes
≤10 weeks' vs >10 weeks' (use all probes)	295 DMRs in total (see supplementary table S4.4)	*SEPTIN4, PRKACB, LGR6, KCNK2, GSDMC, CYP1A2, ABLIM1, SLC26A4, MYLK4, SLC30A4, BDNF-AS, LY86, CTBP2, CD27-AS1, C1R, OIT3, C15orf54, PLD1, ADGRV1, A4GNT, HIF1A, CCR7, THEMIS, CPA3, FGFR2, LINC00880, DDAH1, GLIS3, THEMIS, MIR100HG, IMMP2L, LYPLAL1, SEMA3A, WDR25, SLC37A2, KCNJ1, FAM114A1, MIR548G, DNASE2B, TMEM252, LINC00670, MIR548F1, GUCA1C, DCN, SOX6*
≤10 weeks' vs >10 weeks' (use probes in TSS2000)	*LCP1, KCNK2, CYP1A2, DNMT3A, SYT1, DNASE2B, SLC41A2, SLC30A4, MYLK4, LY86, PPP2R2B, EXOC1L, B3GNT6, CD27-AS1, C5orf64, EPB42, C15orf54, A4GNT, OIT3, CCR7, CPA3, TMEM136, THEMIS, LINC00880, RPTN, CALHM6, GIHCG, MYL1, LIPA, OSBPL1A, GLIS3, TNN, LYST, MGC27382, DRC1, TMEM252, MSTN, PITX2, LINC00670, MIR548F1, MIR101-1, C1QTNF7, GUCA1C, DCN, MLLT6*	*KCNK2, CYP1A2, MYLK4, LY86, CD27-AS1, C15orf54, A4GNT, CCR7, CPA3, THEMIS, LINC00880, GLIS3, TMEM252, MIR548F1, C1QTNF7, GUCA1C, DCN*
≤10 weeks' vs >10 weeks' (use promoter related probes)	*LCP1, MED15, SLC2A3, HIVEP2, PRRC2B, CCR7, TTC39C*	*SLC2A3, CCR7*

Table S4.8 Ninety-one probes that consist gestational age clock.

Probe	**Coefficient**	**Chromosome**	**Position**	**Correlation with GA**	**Correlation p value**	**Correlation FDR**
(Intercept)	-2.797158	NA	NA	NA	NA	NA
cg03264550	0.00156891	chr20	57465448	0.67622036	5.00E-18	1.80E-16
cg07080244	0.00229858	chr11	120051676	0.83159432	3.31E-33	2.75E-31
cg08502383	-0.0023289	chr19	1377233	-0.6116469	3.54E-14	9.19E-13
cg03907570	-0.0036404	chr7	731254	-0.795264	1.66E-28	1.16E-26
cg26490274	0.00517381	chr16	4385435	0.87312229	3.43E-40	3.08E-38
cg24465224	0.00673441	chr8	144984599	0.80966134	3.01E-30	2.35E-28
cg12565585	-0.003446	chr8	105235943	-0.4441823	2.12E-07	1.06E-06
cg12571629	0.01086403	chr2	445430	0.57275405	2.97E-12	5.94E-11
cg00989249	0.03304928	chr5	177548398	0.50697168	1.61E-09	1.77E-08
cg04885749	-0.003433	chr12	64855513	-0.5858128	7.16E-13	1.59E-11
cg21960624	-0.0306869	chr19	35085474	-0.5652266	6.55E-12	1.11E-10

cg04785944	-0.005574	chr9	12554597	-0.6764401	4.83E-18	1.79E-16
cg05527113	0.04438768	chr15	65204428	0.7932288	2.86E-28	1.91E-26
cg03682252	0.00619714	chr6	158401952	0.81442076	7.42E-31	5.86E-29
cg06089960	-3.83E-04	chr1	1498081	-0.7194154	3.37E-21	1.58E-19
cg07920074	0.02798922	chr1	220981523	0.58615345	6.90E-13	1.59E-11
cg08737640	0.00360761	chr17	9683136	0.77250562	5.10E-26	3.06E-24
cg07759237	0.00404935	chr1	150121302	0.77682412	1.81E-26	1.11E-24
cg17378265	0.01691217	chr10	122740411	0.70728387	2.99E-20	1.26E-18
cg00862408	0.0029677	chr11	67177034	0.57387786	2.63E-12	5.53E-11
cg01184449	0.00460536	chr2	42396170	0.64502961	4.71E-16	1.37E-14
cg15085109	0.00223529	chr11	67818764	0.22258011	0.01259922	0.02519844
cg08778805	0.00159499	chr11	92572717	0.85308337	1.49E-36	1.31E-34
cg00405713	0.00803915	chr2	172886349	0.71086644	1.59E-20	6.99E-19
cg11962355	-0.0172894	chr2	182639419	-0.7928197	3.18E-28	2.10E-26
cg01575652	0.00412401	chr22	38077606	0.77812537	1.32E-26	8.19E-25
cg21806580	-0.002609	chr5	18746010	-0.8761593	8.50E-41	7.74E-39
cg14960290	-0.0167713	chr19	12540901	-0.6757342	5.39E-18	1.89E-16
cg03058163	-0.0134449	chr1	1153338	-0.7594767	1.01E-24	5.87E-23
cg21581865	-0.0129632	chr1	32782332	-0.5725602	3.03E-12	5.94E-11
cg04190262	-0.0196783	chr16	69442518	-0.8451306	2.96E-35	2.52E-33
cg00403963	0.02465966	chr19	7250751	0.86580834	8.53E-39	7.59E-37
cg07581823	0.00546007	chr1	155252506	0.72130664	2.37E-21	1.16E-19
cg01501279	-0.0037678	chr6	10434497	-0.5946168	2.65E-13	6.35E-12
cg01584939	0.02749233	chr22	37254322	0.76066258	7.77E-25	4.59E-23
cg05066410	0.01214117	chr3	184802788	0.73400192	2.08E-22	1.13E-20
cg22144260	-0.0047266	chr2	150267714	-0.7171348	5.13E-21	2.36E-19
cg09311336	0.00321293	chr22	38653322	0.52139148	4.54E-10	5.44E-09
cg17358562	0.01725217	chr12	10306335	0.66956164	1.38E-17	4.70E-16
cg25963539	0.00311218	chr16	89181941	0.70552665	4.07E-20	1.67E-18

cg04869116	-0.0957839	chr6	34759164	-0.4946317	4.54E-09	4.46E-08
cg05195559	-0.0384963	chr2	187383871	-0.7469704	1.50E-23	8.39E-22
cg09659734	0.02715526	chr16	56227065	0.31553451	3.38E-04	0.00135163
cg22079271	0.00446427	chr2	233347454	0.61528529	2.26E-14	6.11E-13
cg05030553	-0.0029584	chr21	44061403	-0.6229006	8.74E-15	2.45E-13
cg17965492	0.02578238	chr2	237538965	0.56231293	8.86E-12	1.42E-10
cg08990402	-5.91E-04	chr8	54229644	-0.6620932	4.19E-17	1.34E-15
cg07592012	0.01462176	chr17	79977307	0.70236492	7.04E-20	2.81E-18
cg05243267	-0.0074866	chr4	145275499	-0.7286383	5.92E-22	3.08E-20
cg08047192	0.00516508	chr16	49777781	0.79852638	6.90E-29	4.90E-27
cg06927540	-0.0016131	chr6	155967744	-0.4787928	1.62E-08	1.30E-07
cg00796708	-0.0032156	chr14	88491739	-0.5281127	2.46E-10	3.45E-09
cg09406387	-0.0050103	chr5	58592636	-0.7276096	7.22E-22	3.68E-20
cg01188078	-0.0191578	chr19	42344665	-0.4490407	1.50E-07	9.00E-07
cg02570932	-0.0038815	chr2	160945149	-0.4948516	4.46E-09	4.46E-08
cg03405785	-0.0045232	chr5	148869379	-0.7950494	1.76E-28	1.22E-26
cg03841081	-0.0295745	chr3	150967799	-0.477623	1.78E-08	1.30E-07
cg25468120	-0.0030687	chr3	27259020	-0.6980498	1.47E-19	5.58E-18
cg08253363	-0.0044623	chr19	16529096	-0.8077926	5.17E-30	3.93E-28
cg15930310	3.81E-04	chr1	24517831	0.82625215	1.90E-32	1.56E-30
cg16263857	0.01861462	chr19	10201790	0.85191572	2.34E-36	2.04E-34
cg05375994	-0.0242435	chr16	2960660	-0.7329004	2.59E-22	1.37E-20
cg10030162	0.01280665	chr16	49591203	0.80454885	1.30E-29	9.37E-28
cg01332072	-0.0119277	chr22	39679578	-0.8227036	5.87E-32	4.75E-30
cg01298991	0.02840104	chr2	20422849	0.70777587	2.75E-20	1.18E-18
cg05742691	-0.0126233	chr5	170288255	-0.5986439	1.66E-13	4.16E-12
cg06116549	-0.0062014	chr20	45963043	-0.7860601	1.83E-27	1.17E-25
cg05321946	0.0118018	chr7	55505995	0.65639096	9.57E-17	2.97E-15
cg07294295	0.010116	chr12	3103790	0.11309757	0.20918866	0.20918866

cg15519279	-0.009284	chr7	44635924	-0.7001087	1.04E-19	4.04E-18
cg09798252	-6.47E-04	chr8	59763565	-0.5331452	1.55E-10	2.32E-09
cg11052668	0.00997469	chr6	28626564	0.52741808	2.62E-10	3.45E-09
cg00530710	0.01046899	chr16	49697138	0.75500561	2.70E-24	1.54E-22
cg01201295	0.00117086	chr14	96710013	0.27917169	0.00161708	0.00485125
cg02570063	2.03E-04	chr1	201432127	0.807443	5.72E-30	4.29E-28
cg05983405	8.88E-04	chr10	72360348	0.72091799	2.55E-21	1.22E-19
cg15338665	-0.035624	chr14	81916796	-0.8368822	5.52E-34	4.64E-32
cg04124281	5.84E-04	chr15	80190137	0.66860394	1.60E-17	5.26E-16
cg02016305	0.01036021	chr16	85368682	0.84614232	2.04E-35	1.76E-33
cg06607997	0.00938805	chr1	27699132	0.72310815	1.69E-21	8.47E-20
cg17440098	0.00863694	chr11	34871589	0.80959662	3.07E-30	2.36E-28
cg21290745	-0.0239848	chr20	62721573	-0.5698421	4.04E-12	7.28E-11
cg18834254	0.0029833	chr2	241286505	0.73791111	9.58E-23	5.27E-21
cg20111980	0.02807282	chr4	6676796	0.82131361	9.06E-32	7.25E-30
cg07221039	0.00661016	chr20	61423954	0.77924739	1.00E-26	6.33E-25
cg13971552	0.00113103	chr19	10463352	0.80694916	6.58E-30	4.81E-28
cg09308580	-0.0155917	chr2	43405947	-0.6536073	1.42E-16	4.27E-15
cg08218149	-0.009856	chr11	70072100	-0.7143371	8.53E-21	3.84E-19
cg14583126	1.47E-04	chr5	66452973	0.79499092	1.79E-28	1.22E-26
cg05904364	-0.0140176	chr2	20650792	-0.8070499	6.40E-30	4.73E-28
cg11484828	-0.0156183	chr22	43139880	-0.7879406	1.13E-27	7.37E-26

References

1. Schroeder DI, Blair JD, Lott P, Yu HOK, Hong D, Crary F, et al. The human placenta methylome. Proceedings of the national academy of sciences. 2013;110(15):6037-42.
2. Gude NM, Roberts CT, Kalionis B, King RG. Growth and function of the normal human placenta. Thrombosis research. 2004;114(5-6):397-407.
3. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet. 2016 Aug;17(8):487-500. DOI: 10.1038/nrg.2016.59.
4. Jang HS, Shin WJ, Lee JE, Do JT. CpG and non-CpG methylation in epigenetic gene regulation and brain function. Genes. 2017;8(6):148.
5. Reddington JP, Pennings S, Meehan RR. Non-canonical functions of the DNA methylome in gene regulation. Biochemical Journal. 2013;451(1):13-23.
6. Goldberg AD, Allis CD, Bernstein E. Epigenetics: a landscape takes shape. Cell. 2007 Feb 23;128(4):635-8. DOI: 10.1016/j.cell.2007.02.006.
7. Wilson SL, Robinson WP. Utility of DNA methylation to assess placental health. Placenta. 2018;64:S23-S8.
8. Bianco-Miotto T, Mayne BT, Buckberry S, Breen J, Lopez CMR, Roberts CT. Recent progress towards understanding the role of DNA methylation in human placental development. Reproduction (Cambridge, England). 2016;152(1):R23-R30.
9. Novakovic B, Yuen RK, Gordon L, Penaherrera MS, Sharkey A, Moffett A, et al. Evidence for widespread changes in promoter methylation profile in human placenta in response to increasing gestational age and environmental/stochastic factors. BMC genomics. 2011;12(1):529.

10. Nordor AV, Nehar-Belaid D, Richon S, Klatzmann D, Bellet D, Dangles-Marie V, et al. The early pregnancy placenta foreshadows DNA methylation alterations of solid tumors. Epigenetics. 2017;12(9):793-803.

11. Jauniaux E, Watson AL, Hempstock J, Bao Y-P, Skepper JN, Burton GJ. Onset of maternal arterial blood flow and placental oxidative stress: a possible factor in human early pregnancy failure. The American journal of pathology. 2000;157(6):2111-22.

12. Miller S, Dykes D, Polesky H. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic acids research. 1988;16(3):1215.

13. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014 May 15;30(10):1363-9. DOI: 10.1093/bioinformatics/btu049.

14. Konishi T. Principal component analysis for designed experiments. BMC Bioinformatics. 2015;16 Suppl 18:S7. DOI: 10.1186/1471-2105-16-S18-S7.

15. Wan Q, Leemaqz SY-L, Pederson SM, McCullough D, McAninch DC, Jankovic-Karasoulos T, et al. Quality control measures for placental sample purity in DNA methylation array analyses. Placenta. 2019;88:8-11.

16. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome biology. 2016;17(1):208.

17. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. Nucleic acids research. 2015;44(3):e20-e.

18. Triche Jr TJ, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of illumina infinium dna methylation beadarrays. Nucleic acids research. 2013;41(7):e90-e.

19. Xu Z, Langie SA, De Boever P, Taylor JA, Niu L. RELIC: a novel dye-bias correction method for Illumina Methylation BeadChip. BMC genomics. 2017;18(1):4.

20. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. Epigenomics. 2011;3(6):771-84.

21. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2012;29(2):189-96.

22. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC bioinformatics. 2010;11(1):587.

23. Wahl S, Fenske N, Zeilinger S, Suhre K, Gieger C, Waldenberger M, et al. On the potential of models for location and scale for genome-wide DNA methylation data. BMC bioinformatics. 2014;15(1):232.

24. Horvath S. DNA methylation age of human tissues and cell types. Genome biology. 2013;14(10):3156.

25. Ritchie ME, Diyagama D, Neilson J, van Laar R, Dobrovic A, Holloway A, et al. Empirical array quality weights in the analysis of microarray data. BMC bioinformatics. 2006;7(1):261.

26. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, Lord RV, et al. De novo identification of differentially methylated regions in the human genome. Epigenetics & chromatin. 2015;8(1):6.

27. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011;98(4):288-95.

28. Eckmann-Scholz C, Bens S, Kolarova J, Schneppenheim S, Caliebe A, Heidemann S, et al. DNA-methylation profiling of fetal tissues reveals marked epigenetic differences between chorionic and amniotic samples. PLoS One. 2012;7(6):e39014.

29. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics. 2015;32(2):286-8.

30. Andrews S. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom; 2010.

31. Schubert M, Lindgreen S, Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC research notes. 2016;9(1):88.

32. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.

33. Institute B. "Picard Tools." GitHub repository. 2018.

34. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30.

35. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome biology. 2010;11(3):R25.

36. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40.

37. Kuhn M. Building Predictive Models in R Using the caret Package. Journal of Statistical Software. 2008 Nov;28(5):1-26.

38. Court F, Tayama C, Romanelli V, Martin-Trujillo A, Iglesias-Platas I, Okamura K, et al. Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. Genome Res. 2014 Apr;24(4):554-69. DOI: 10.1101/gr.164913.113.

39. Wan Q, Leemaqz SY-L, Pederson SM, McCullough D, McAninch DC, Jankovic-Karasoulos T, et al. Quality control measures for placental sample purity in DNA methylation array analyses. Placenta. 2019.

40. Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome biology. 2015;16(1):22.

41. Lizio M, Abugessaisa I, Noguchi S, Kondo A, Hasegawa A, Hon CC, et al. Update of the FANTOM web resource: expansion to provide additional transcriptome atlases. Nucleic acids research. 2019;47(D1):D752-D8.

42. Jeziorska DM, Murray RJ, De Gobbi M, Gaentzsch R, Garrick D, Ayyub H, et al. DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. Proceedings of the National Academy of Sciences. 2017;114(36):E7526-E35.

43. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics. 2017;33(24):3982-4.

44. Avila L, Yuen R, Diego-Alvarez D, Penaherrera M, Jiang R, Robinson W. Evaluating DNA methylation and gene expression variability in the human term placenta. Placenta. 2010;31(12):1070-7.

45. Cardenas A, Gagné-Ouellet V, Allard C, Brisson D, Perron P, Bouchard L, et al. Placental DNA methylation adaptation to maternal glycemic response in pregnancy. Diabetes. 2018;67(8):1673-83.

46. Robinson WP, Price EM. The human placental methylome. Cold Spring Harbor perspectives in medicine. 2015;5(5):a023044.

47. Chuong EB, Rumi MK, Soares MJ, Baker JC. Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nature genetics. 2013;45(3):325.

48. Novakovic B, Saffery R. Placental pseudo-malignancy from a DNA methylation perspective: unanswered questions and future directions. Frontiers in genetics. 2013;4.

49. Trusolino L, Comoglio PM. Scatter-factor and semaphorin receptors: cell signalling for invasive growth. Nature Reviews Cancer. 2002;2(4):289.

50. Gavin MA, Rasmussen JP, Fontenot JD, Vasta V, Manganiello VC, Beavo JA, et al. Foxp3-dependent programme of regulatory T-cell differentiation. Nature. 2007;445(7129):771.

51. Barbaux S, Gascoin-Lachambre G, Buffat C, Monnier P, Mondon F, Tonanny M-B, et al. A genome-wide approach reveals novel imprinted genes expressed in the human placenta. Epigenetics. 2012;7(9):1079-90.

52. Li Z, Liu C, Xie Z, Song P, Zhao RC, Guo L, et al. Epigenetic dysregulation in mesenchymal stem cell aging and spontaneous differentiation. PloS one. 2011;6(6):e20526.

53. Smallwood SA, Kelsey G. De novo DNA methylation: a germ cell perspective. Trends in Genetics. 2012;28(1):33-42.

54. He X-J, Chen T, Zhu J-K. Regulation and function of DNA methylation in plants and animals. Cell research. 2011;21(3):442.

55. Suryawanshi H, Morozov P, Straus A, Sahasrabudhe N, Max KE, Garzia A, et al. A single-cell survey of the human first-trimester placenta and decidua. Science Advances. 2018;4(10):eaau4788.

56. Red-Horse K, Kapidzic M, Zhou Y, Feng K-T, Singh H, Fisher SJ. EPHB4 regulates chemokine-evoked trophoblast responses: a mechanism for incorporating the human placenta into the maternal circulation. Development. 2005;132(18):4097-106.

57. Langdown ML, Sugden MC. Enhanced placental GLUT1 and GLUT3 expression in dexamethasone-induced fetal growth retardation. Molecular and Cellular Endocrinology. 2001;185(1-2):109-17.

58. Jeckel K, Boyarko A, Bouma G, Winger Q, Anthony R. Chorionic somatomammotropin impacts early fetal growth and placental gene expression. Journal of Endocrinology. 2018;237(3):301-10.

59. Komatsu N, Matsueda S, Tashiro K, Ioji T, Shichijo S, Noguchi M, et al. Gene expression profiles in peripheral blood as a biomarker in cancer patients receiving peptide vaccination. Cancer. 2012;118(12):3208-21.

60. Starnes T, Broxmeyer HE, Robertson MJ, Hromas R. Cutting edge: IL-17D, a novel member of the IL-17 family, stimulates cytokine production and inhibits hemopoiesis. The Journal of Immunology. 2002;169(2):642-6.

61. O'Sullivan T, Saddawi-Konefka R, Gross E, Tran M, Mayfield SP, Ikeda H, et al. Interleukin-17D mediates tumor rejection through recruitment of natural killer cells. Cell reports. 2014;7(4):989-98.

62. Hatt L, Aagaard MM, Graakjaer J, Bach C, Sommer S, Agerholm IE, et al. Microarray-based analysis of methylation status of CpGs in placental DNA and maternal blood DNA–potential new epigenetic biomarkers for cell free fetal DNA-based diagnosis. PLoS One. 2015;10(7).

63. Mayne BT, Leemaqz SY, Smith AK, Breen J, Roberts CT, Bianco-Miotto T. Accelerated placental aging in early onset preeclampsia pregnancies identified by DNA methylation. Epigenomics. 2017 Mar;9(3):279-89. DOI: 10.2217/epi-2016-0103.

64. Lee Y, Choufani S, Weksberg R, Wilson SL, Yuan V, Burt A, et al. Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels. Aging (Albany NY). 2019;11(12):4238.

65. McEwen LM, Jones MJ, Lin DTS, Edgar RD, Husquin LT, MacIsaac JL, et al. Systematic evaluation of DNA methylation age estimation with common preprocessing methods and the Infinium MethylationEPIC BeadChip array. Clinical epigenetics. 2018;10(1):1-9.

66. Darst BF, Malecki KC, Engelman CD. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. BMC genetics. 2018;19(1):65.

67. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018;10(4):573.

Statement of Authorship

Title of Paper	DNA methylation profiling of human placenta samples across early gestation.
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Qianhui Wan		
Contribution to the Paper	Performed the data analyses and wrote the manuscript.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	23/03/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	James Breen		
Contribution to the Paper	Conceived the study design, supervised the work, assisted in coding and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Shalem Yiner-Lee Leemaqz		
Contribution to the Paper	Assisted with statistics and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Stephen Martin Pederson		
Contribution to the Paper	Assisted with statistics and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Tanja Jankovic-Karasoulos		
Contribution to the Paper	Laboratory work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Dale Christopher McAninch		
Contribution to the Paper	Laboratory work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Dylan McCullough		
Contribution to the Paper	Sample collection and laboratory work.		
Signature		Date	23/03/20

Name of Co-Author	Melanie Denise Smith		
Contribution to the Paper	Assisted in coding and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Konstantinos Justinian Bogias		
Contribution to the Paper	Assisted in coding and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Ning Liu		
Contribution to the Paper	Assisted in coding.		
Signature		Date	23/03/20

Name of Co-Author	Callista Leonie Mulder		
Contribution to the Paper	Edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Claire Trelford Roberts		
Contribution to the Paper	Conceived the study design, supervised the work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Tina Bianco-Miotto		
Contribution to the Paper	Conceived the study design, supervised the work and edited the manuscript.		
Signature		Date	23/03/20

Please cut and paste additional co-author panels here as required.

5 DNA methylation profile of maternal peripheral leukocytes across early to mid-pregnancy

Abstract

DNA methylation has been shown repeatedly to be a suitable biomarker for early prediction of health outcomes that can potentially be widely used in a clinical setting [1]. There are some studies that attempt to use DNA methylation in circulating maternal leukocytes to predict pregnancy complications but to do this, we need to define how DNA methylation changes throughout uncomplicated pregnancies. Therefore, the aim of this study was to profile peripheral leukocyte methylomes from 131 pregnant women across 6-23 weeks' gestation using Illumina Infinium MethylationEPIC BeadChips. We estimated cell proportion for samples and found these also changed across early gestation with percentage of neutrophils increased and other cell types decreased across early gestation. DNA methylation changes related to maternal smoking, maternal age and gestational age were also identified. The predicted DNA methylation age of pregnant women was higher than the chronological age indicating pregnancy may accelerate aging. To our knowledge, this is the first reported study using 131 maternal leukocyte-containing buffy coat samples with 6-23 weeks of pregnancy to reveal the DNA methylation changes associated with fetal gestational age. This study will lay the foundation for real time assessment of pregnancy health using DNA methylation in maternal circulating leukocytes.

Keywords: maternal leukocytes, DNA methylation, early gestation, early pregnancy

5.1 Introduction

Epigeneticists study altered phenotypes caused by changes in molecular groups attached to chromatin without changes in DNA sequences [2]. Epigenetic mechanisms include DNA methylation, histone modification and noncoding RNAs [3]. DNA methylation is the first identified and best studied epigenetic modification in mammals [4]. By adding a methyl group at 5-cytosine (5mC), DNA methylation of CpG islands at gene promoters are associated with gene regulation and have been widely studied. The 5mCs at gene promoters can potentially alter chromatin structure, block transcription initiation and repress gene expression [5].

Fetal development is associated with changes of DNA methylation in multiple fetal tissues such as placenta and cord blood but also in maternal leukocytes [6]. Studies using samples from healthy and complicated pregnancies have shown that pregnancy complications are accompanied by DNA methylation changes in placenta, cord blood and maternal blood [7-9]. DNA methylation also provides an important link between the genome and maternal environmental exposures such as smoking during pregnancy and is an important part of translational epigenetics [10]. Compared with

studies on pregnancy complications and environmental adverse effects, there are fewer that have focused on identifying DNA methylation changes in normal pregnancies across gestation which would provide an important baseline.

DNA methylation in peripheral leukocytes is widely studied since it is the most commonly available tissue that can be collected non-invasively for use in pregnancy [11]. We hypothesise that there are DNA methylation differences between maternal leukocyte samples of different clinical phenotypes and also difference between leukocytes from pregnant and non-pregnant women. Up to now, there has only been one study focusing on identifying the DNA methylation difference between blood samples from pregnant and non-pregnant women using 27K arrays [12]. In the present study, we have used Illumina Infinium MethylationEPIC (EPIC) BeadChips to identify changes in DNA methylation profiles of maternal leukocytes across early gestation which could be considered as biomarkers that could potentially be used to monitor fetal development.

5.2 Materials and Methods

5.2.1 Ethics statement

Same as the ethics statement in Chapter 4.

5.2.2 Sample Description

Seventy first trimester (6-13 weeks') and sixty-one second trimester (14-23 weeks') maternal blood samples were obtained from women with age ranges from 16-45 years old and undergoing elective pregnancy terminations at the Pregnancy Advisory Centre in Woodville, South Australia. Gestational age was determined using transabdominal ultrasonography. Females with infection, endocrine abnormalities, antiphospholipid syndrome or other known complications were excluded from the study. At the time of termination, 9 mL of whole blood was collected in an EDTA blood collection tube. Blood samples were separated by centrifugation followed by the removal of the plasma. The buffy coat layer, which contains all the nucleated cells (leukocytes) in blood, was isolated and washed in TE buffer before being stored at -80°C until DNA was extracted. A modified version of the TES protocol [13] was used to extract DNA. For each sample, 1 µg of DNA was sent to PathWest Laboratory Medicine (QEII Medical Centre, Perth, Western Australia) for bisulfite-conversion and hybridisation to the Illumina Infinium® MethylationEPIC BeadChips according to manufacturer's instructions.

Two data sets, GSE74738 [14] and GSE123914 [15], from GEO database were also used in this study. Maternal leukocyte samples from data set GSE74738 are samples obtained from pregnant women, and leukocyte samples from data set GSE123914 are

from non-pregnant women with age ranges from 51-64 years old. The platforms and number of samples used in this study were listed in Table 5.1.

Table 5.1 Samples used in this study

Trimester	Maternal leukocyte samples (our study)	Maternal leukocyte samples (GSE74738)	Non-pregnant female leukocytes samples (GSE123914)
Array type	EPIC	450K	EPIC
Total number	131	10	69

5.2.3 EPIC array data analysis

5.2.3.1 Quality control

Same as the quality control method in Chapter 4.

5.2.3.2 Probes on microarray and filtering of failed probes

Illumina Infinium® MethylationEPIC array contains 866238 probes [16]. Failed and unwanted probes were filtered. The failed probes removed included 7732 probes with detection P > 0.01, 13490 probes with probes < 3 beads in 5% of the 131 samples, 25836 probes with SNPs at CpG/SBE sites on probes [17], and 42249 cross-reactive probes [16]. In favour of downstream analysis, 133603 probes with SNPs at probe body (not located at CpG/SBE sites) were also removed to decrease the disturbance

variables. In total, 222910 probes were removed, and 643328 probes remained for downstream analyses. Details of maternal leukocyte samples and filtered probes are listed in Supplementary Figure S5.1.

5.2.3.3 Background and dye bias correction

Same as the method in Chapter 4.

5.2.3.4 Data normalisation

Same as the method in Chapter 4 and samples from different studies were normalised together with BMIQ method.

5.2.3.5 Identification of differentially methylated positions (DMPs)

Similar as method for detecting DMPs in Chapter 4, M values were used for analysing differentially methylated positions of maternal leukocytes across early pregnancy. Batch effects were adjusted in the linear models and relative quality weights for samples were estimate based on sample variances in groups with highly variable samples [18] down-weighted in the models [19]. The equation of the model is $\mathrm{DNAm} =$

$\beta_1 Trimester + \beta_2 Batch + \sum_{i=3}^{9} \beta_3 \, CellType_i + \epsilon, \epsilon \sim N(\mu, \sigma^2)$, where $CellType_i$ is the proportion of ith cells (CD8T, CD4T, NK, B cell, Mono and Gran) [20]. DNAm is the M value of each probe in each sample, $Trimester$ is the group we are interested, $Batch$ represents the array received from three different batches. Differentially methylated positions were identified between groups of samples from 6-13 weeks (first trimester) of pregnancy and 14-23 weeks (second trimester) of gestation with false discovery rate (FDR) less than 0.05 and the change of beta values ($|\Delta\beta| > 0.2$) greater than 0.2 between groups.

5.2.3.6 Maternal age analysis

Normalised beta values of 131 samples were transformed to M values for maternal age analysis. Horvath epigenetic clock implemented in R package *wateRmelon* [21] was used for maternal age prediction. The accuracy of predicted age was determined by the correlation coefficients between DNA methylation age and maternal age.

5.2.3.7 Estimation of cell-type composition

Cell-type composition of each maternal leukocyte sample was estimated with the reference-based method first published by Houseman et al. [22], which uses DNA methylation reference profiles of individual cell-types to estimate the cell-type

composition of each sample. The proportions of the four major lymphocyte subtypes (i.e. CD4+ T, CD8+ T, CD19+ B and CD56+ NK cells) and two myeloid cells types (i.e. monocytes, neutrophils) were estimated. This was done using the R package *minfi* [17], because this package enables the estimation of cell proportions of leukocyte samples from Illumina Infinium HumanMethylation450 BeadChip platform and the method used in this package is based on published methylation data of flow-sorted cells [23] using algorithms developed by Houseman et al. [22]. The estimated cell type proportions of maternal leukocytes were adjusted in the differential methylation analyses [24].

5.2.3.8 GO and KEGG analyses

The gene ontology (GO), KEGG pathway analysis and gene set analysis (GSA) were conducted using the *missMethyl* [25] R package to determine if differential methylation of genes in particular pathways were associated with differences between sample groups. The *missMethyl* package was developed specifically for the Illumina HumanMethylation450K and EPIC array, since *gometh* and *gsameth* functions in this package [25] tested for gene ontology and gene set enrichment for significant positions, taking into account the different number of probes for each gene on the array.

5.2.3.9 Annotation

Same as the method in Chapter 4.

5.3 Results

5.3.1 Maternal blood DNA methylation profiles across early gestation

PCA and density plots show that the 131 maternal leukocyte samples from pregnant women at 6-23 weeks' gestation (meta data in Supplementary Table S5.1) cluster together. Maternal blood samples from our study clustered with publicly available maternal blood samples from pregnant women but not with samples from non-pregnant women (Figure 5.1A). When just EPIC DNA methylation profiles from blood samples from pregnant women and non-pregnant women were assessed, they clustered separately (Figure 5.1B). The density plot showed, as expected, that there were more methylated sites than unmethylated sites in maternal leukocyte DNA (Figure 5.1C). A PCA plot based on all 131 maternal leukocyte samples from our study showed that the DNA methylation variation of maternal blood is subtle across early to mid-pregnancy and samples did not cluster according to gestational age (Figure 5.1D). Heatmap of PC1 leading probes (n=6433) showed that most of these sites at CpG islands were unmethylated across gestation with slight changes.

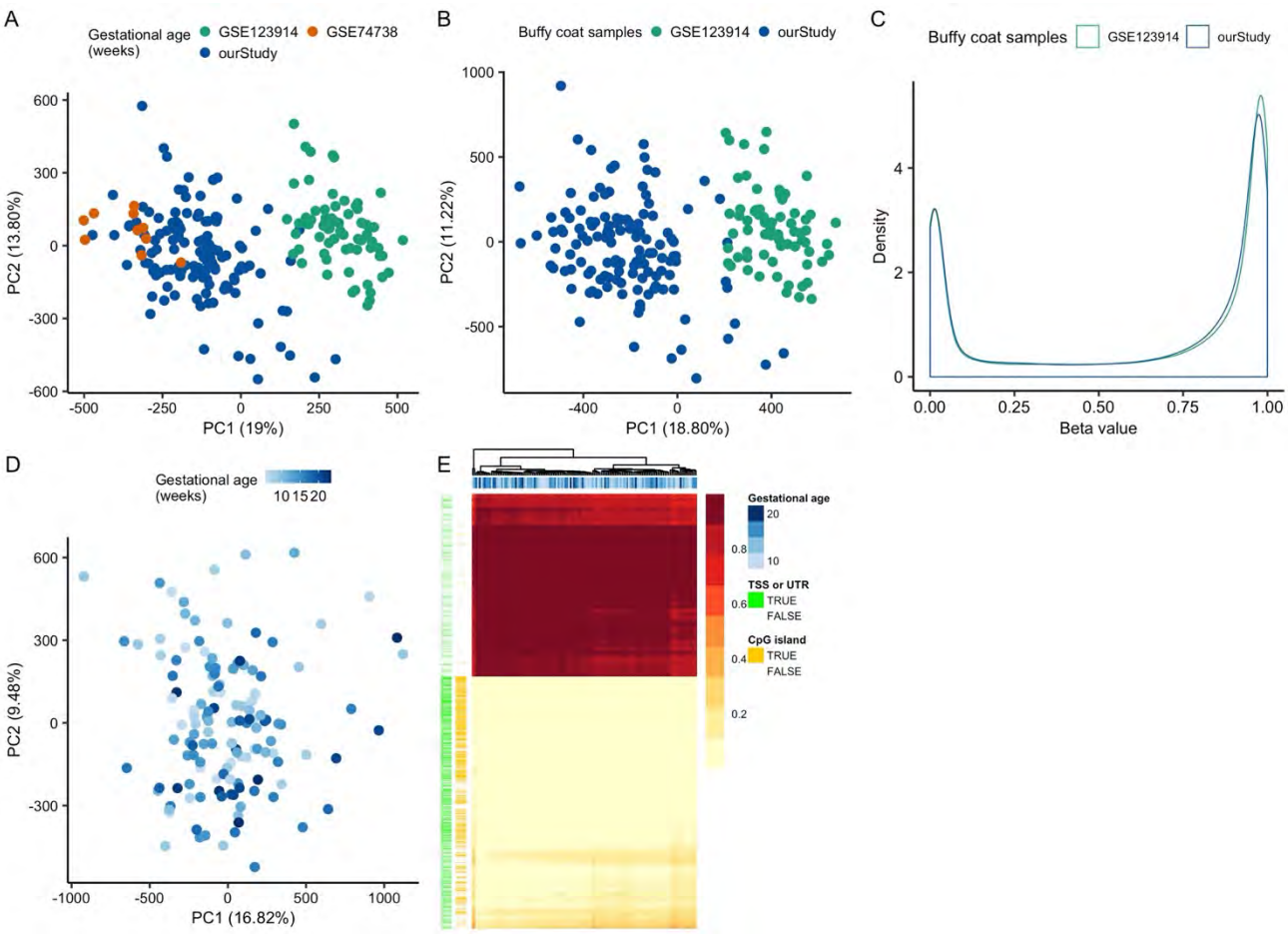

Figure 5.1 Maternal leukocyte DNA methylation profiles across early to mid-pregnancy. PCA plot based on 450K probes (A) and EPIC probes (B) showed that leukocyte samples from pregnant women (GSE74738 and our study) and non-pregnant women (GSE123914) had different DNA methylation profiles. (C) There are more methylated sites than unmethylated sites for leukocyte samples. (D) DNA methylation profiles of maternal leukocytes did not separate due to gestational age in pregnancy. (E) Heatmap shows subtle changes ($|\Delta\beta| < 0.2$) of DNA methylation of PC1 leading probes (top 1%, n=6433) across early to mid-pregnancy.

5.3.2 High variable probes are enriched in immune response

To investigate the variation of maternal peripheral leukocyte DNA methylation across samples from 6-23 weeks' gestation, we analysed the change of DNA methylation for the high variable probes with SD of beta values higher than 0.2 across all samples,

which is also used in previous publications as cutoff for the most variable CpGs [26]. In total, there were 495 probes that passed this threshold. In Figure 5.2A, each row represents a DNA methylation site with colour denoting the position of these sites and each column is a maternal leukocyte sample with blue colour representing the fetal gestational age. The heatmap of these 495 sites showed that even though these sites varied across samples, they did not change across gestational age groups (Figure 5.2A). Unsurprisingly, these probes were enriched in terms related to the immune response (Figure 5.2B) indicating that the DNA methylation differences were between leukocytes across individuals.

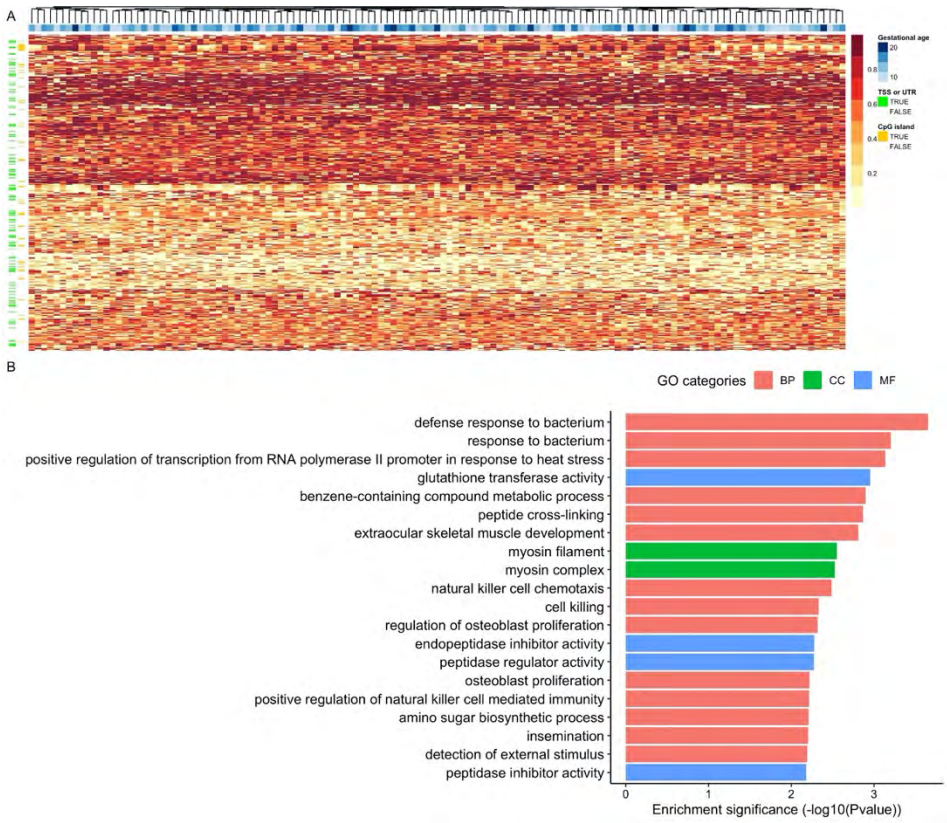

Figure 5.2 High variable probes across maternal leukocyte samples from 6-23 weeks' pregnancy. (A) Heatmap of 495 high variable probes across gestation. (B) Gene ontology analyses of 495 high variable probes (top 20 GO terms ranked by p-values). BP: biological process, MF: molecular function, CC: cellular component.

5.3.3 Cell proportion of maternal leukocytes across early to mid-gestation

Since different cell types including CD4 T cells, CD8 T cells, B cells, NK cells, monocytes and neutrophils have different functions and hence different DNA methylation profiles (need a reference here) they can impact overall DNA methylation patterns of maternal leukocytes, therefore the cell type proportions in each sample were estimated [27]. Using data from our study, the cell proportions of 61 maternal leukocyte samples from second trimester were identified as different compared to the

70 samples from first trimester, with the percentage of neutrophils increasing and the percentage of other cell types decreasing (Figure 5.3A). Since an increase of neutrophils was most noticeable from 13 weeks' gestation onwards (which is also the cut off for first and second trimesters), we divided samples into two groups i.e. samples with gestational age ≤ 13 weeks' and samples with gestational age >13 weeks' (Figure 5.3A). The proportion of four major lymphocyte subtypes (i.e. CD4 T, CD8 T, B and NK cells) and two myeloid cells types (i.e. monocytes, neutrophils) are different between samples from pregnant and non-pregnant women (Figure 5.3B). There was an increase in neutrophils and CD8T cells in blood samples from pregnant women compared with non-pregnant women but a decrease in CD4T, NK, B cells and monocytes.

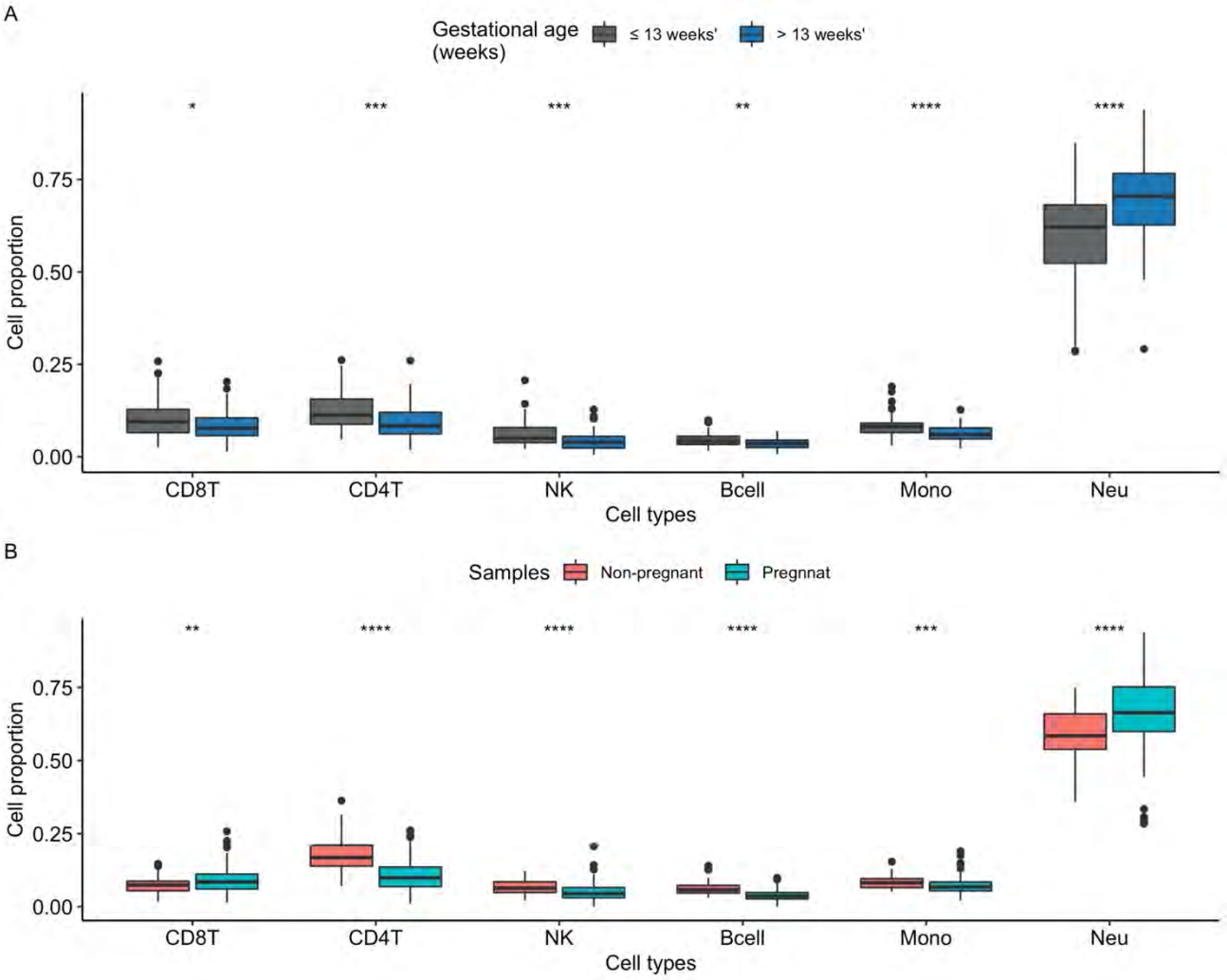

Figure 5.3 Cell proportion may change across early to mid-pregnancy. (A) Estimated cell proportion of maternal leukocytes across early gestation. (B) Estimated cell proportion of leukocytes from non-pregnant and pregnant women.

5.3.4 DNA methylation of maternal peripheral leukocytes associated with gestational age, maternal age and smoking status

To investigate the influences of phenotypes including maternal smoking, maternal age and fetal gestational age on DNA methylation of maternal leukocytes, singular value decomposition (SVD) analyses were performed and the results showed that there are

DNA methylation changes related to gestational age, smoking, maternal age and cell proportion (Figure 5.4A-B). In order to investigate these changes, leading probes in PC1 and PC3 were analysed. We identified 17 sites that were associated with maternal smoking with group mean beta difference > 0.1 and 3 sites were significantly correlated (p-value <0.05) with PC1. DNA methylation of *SLIT3* and *GLIPR2* were significantly correlated with PC1 and DNA methylation of these two sites were increased and decreased, respectively in the smoking group compared with the non-smoking group (Figure 5.4C and Supplementary Table S5.2). Another 13 sites identified from PC1 were significantly correlated with gestational age (FDR <0.05 and correlation >0.2) and they were associated with 9 genes (*GRB10*, *HOXA10-AS*, *RASGRP1*, *LINS*, *SLC25A37*, *NADK*, *JAK3*, *KDM4B* and *IGF2BP2*) (Figure 5.4D and Supplementary Table S5.3). We also identified 59 sites from PC3 that were significantly correlated with maternal age (FDR <0.05 and correlation >0.2) and 22 of these sites co-localized with 5'UTR or transcription start site (TSS) of genes. Two sites at the gene body of *DIRAS* had higher DNA methylation with increasing maternal age (Figure 5.4E and Supplementary Table S5.4).

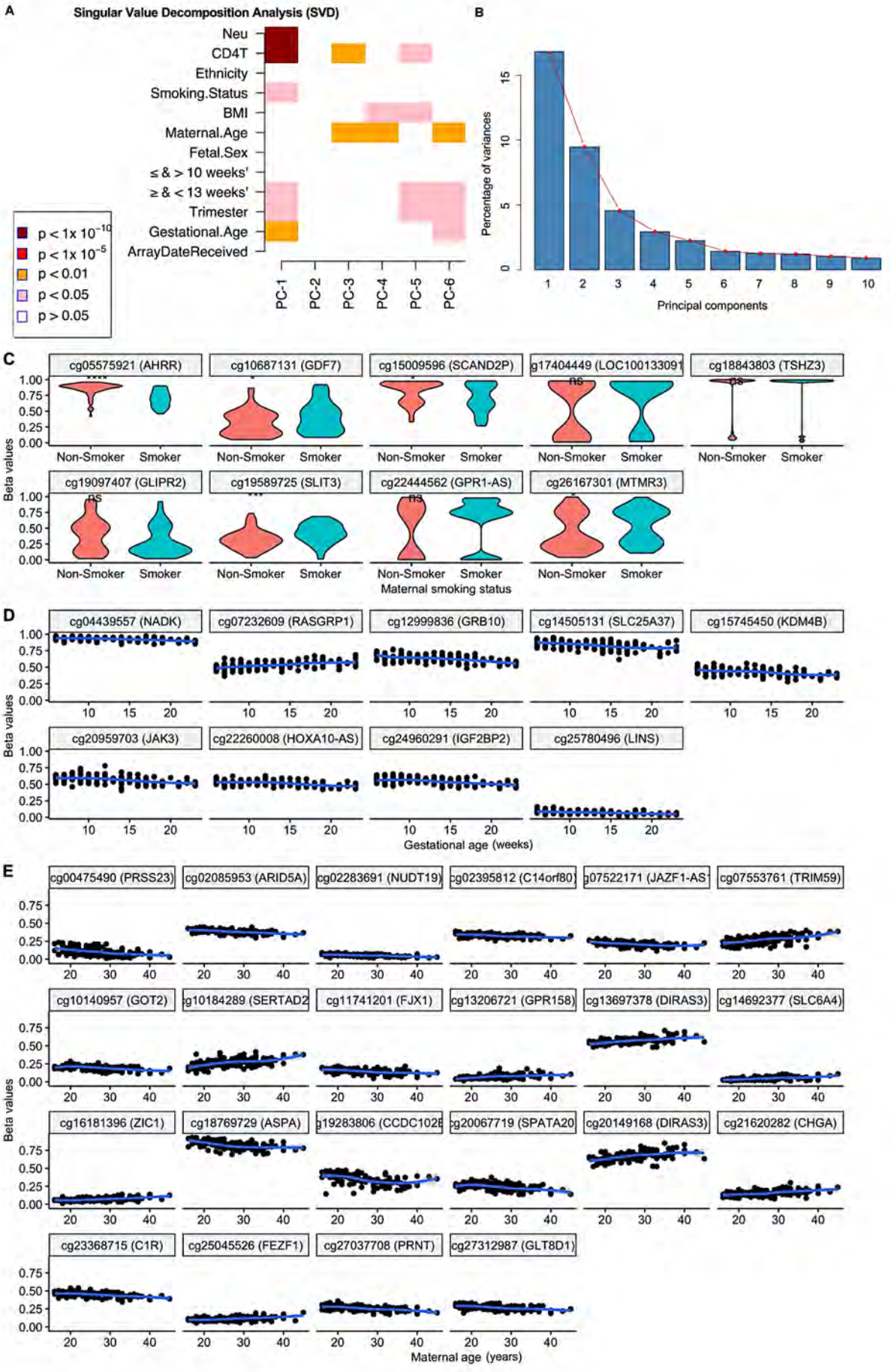
A
Singular Value Decomposition Analysis (SVD)
Neu
CD4T
Ethnicity
Smoking.Status
BMI
Maternal.Age
Fetal.Sex
≤ & > 10 weeks'
≥ & < 13 weeks'
Trimester
Gestational.Age
ArrayDateReceived
PC-1
PC-2
PC-3
PC-4
PC-5
PC-6
p < 1x 10^-10
p < 1x 10^-5
p < 0.01
p < 0.05
p > 0.05
B
Percentage of variances
Principal components
C
cg05575921 (AHRR)
cg10687131 (GDF7)
cg15009596 (SCAND2P)
cg17404449 (LOC100133091)
cg18843803 (TSHZ3)
cg19097407 (GLIPR2)
cg19589725 (SLIT3)
cg22444562 (GPR1-AS)
cg26167301 (MTMR3)
Beta values
Non-Smoker
Smoker
Maternal smoking status
D
cg04439557 (NADK)
cg07232609 (RASGRP1)
cg12999836 (GRB10)
cg14505131 (SLC25A37)
cg15745450 (KDM4B)
cg20959703 (JAK3)
cg22260008 (HOXA10-AS)
cg24960291 (IGF2BP2)
cg25780496 (LINS)
Gestational age (weeks)
E
cg00475490 (PRSS23)
cg02085953 (ARID5A)
cg02283691 (NUDT19)
cg02395812 (C14orf80)
cg07522171 (JAZF1-AS1)
cg07553761 (TRIM59)
cg10140957 (GOT2)
cg10184289 (SERTAD2)
cg11741201 (FJX1)
cg13206721 (GPR158)
cg13697378 (DIRAS3)
cg14692377 (SLC6A4)
cg16181396 (ZIC1)
cg18769729 (ASPA)
cg19283806 (CCDC102B)
cg20067719 (SPATA20)
cg20149168 (DIRAS3)
cg21620282 (CHGA)
cg23368715 (C1R)
cg25045526 (FEZF1)
cg27037708 (PRNT)
cg27312987 (GLT8D1)
Maternal age (years)

Figure 5.4 DNA methylation changes associated with gestational age, smoking status and maternal age. (A) SVD analysis for all the maternal leukocyte samples. Variances in PC1 mainly related to cell types and gestational age and variances in PC3 mainly related to cell types and maternal age. (B) Scree plot for the first 10 principal components. PC1 represents the largest variance in this data set. (C) DNA methylation of sites related with maternal smoking status. (D) DNA methylation of sites related with gestational age. (E) DNA methylation of sites related with maternal age.

5.3.5 DNA methylation is similar in maternal leukocytes before and after 13 weeks' gestation

Since there is an increase in the estimated neutrophil cell proportion in samples > 13 weeks' gestation compared to ≤ 13 weeks' gestation, differential methylation analyses were applied for 131 maternal peripheral leukocyte samples at this gestational age threshold. Since DNA methylation subtly changes across early to mid-pregnancy, there were no positions identified as significantly different between groups with cell composition adjusted in the linear models. We did not identify DMPs using all probes. However, 555 sites in promoter regions with beta changes higher than 0.05 between groups were investigated (Figure 5.5A). Each row of the heatmap (Figure 5.5B) represents a DNA methylation site with colour denoting whether these DNA methylation positions are on CpG islands, transcription start sites (TSS) or UTR regions and each column is a maternal leukocyte sample with blue colour representing the fetal gestational age. The heatmap showed that DNA methylation of these sites

slightly increased or decreased across early to mid-pregnancy and these sites were enriched for T cell activation (Figure 5.5C).

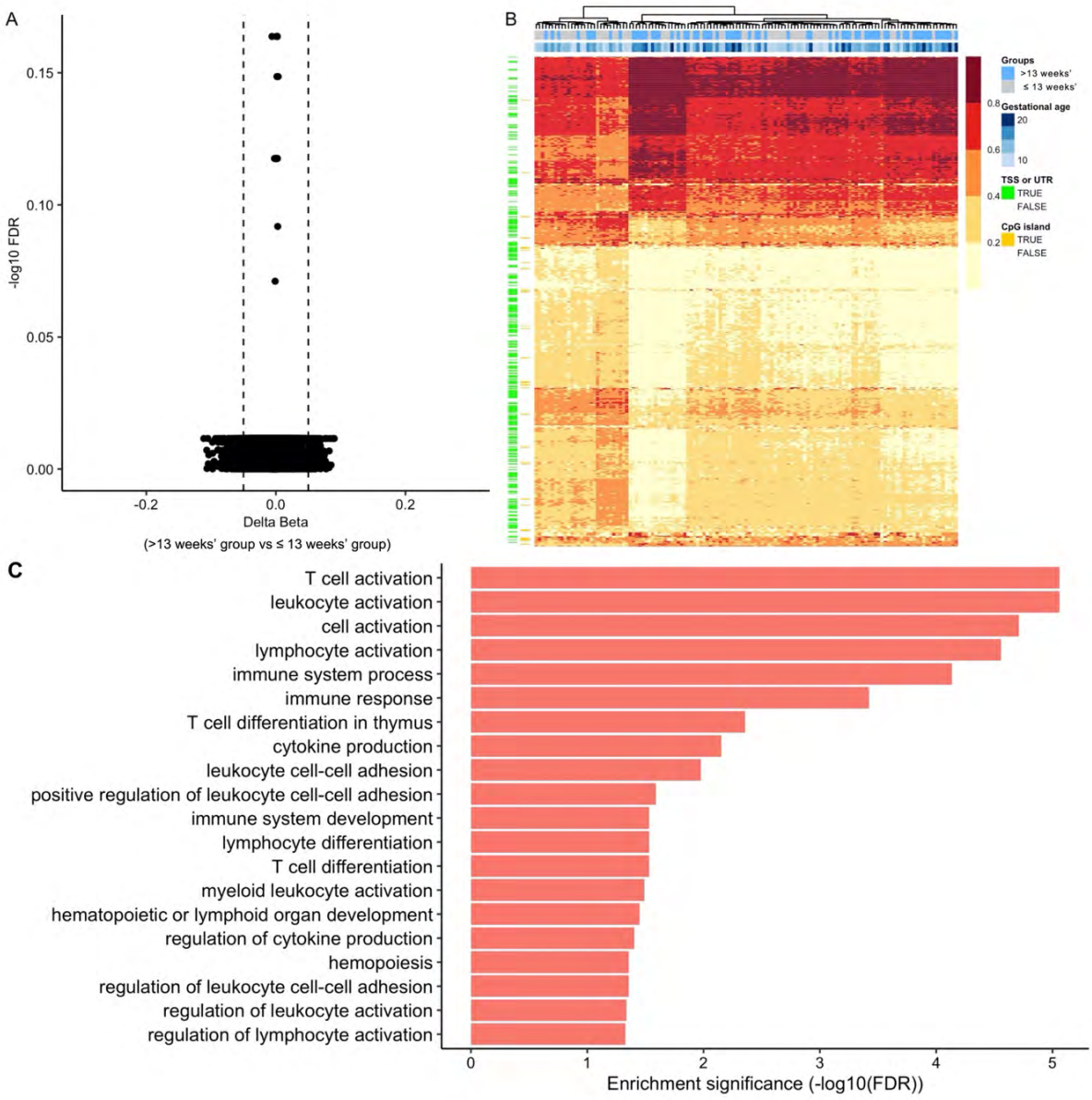

Figure 5.5 Comparison of the DNA methylation profile of maternal peripheral leukocytes up to and including 13 weeks' gestation with after that time. (A) Volcano plot showed that DNA methylation is similar between the two groups. (B) Heatmap of the selected 555 sites that have DNA methylation changes > 0.05. (C) Top 20 terms from GO analyses of the genes colocalized with these 555 sites (FDR<0.05).

5.3.6 Prediction of age

Since the epigenetic clock has repeatedly been shown to be accurate in estimating age [18], in this study, leukocyte samples were used to predict maternal age. Using the published epigenetic clock for blood samples based on 353 DNA methylation sites, here we show that DNA methylation profile of maternal peripheral leukocytes can be used for estimating maternal age. The estimated maternal ages for pregnant women were overall higher than the real maternal age with a root mean squared error (RMSE) of 9.1 years (Figure 5.6A). The estimated age of non-pregnant women had less variation than for pregnant women with a RMSE of 3.9 years, however, overall, the correlation between maternal age and the DNA methylation age is not high and this could be due to the variation between samples which could be caused by maternal age and other unknown phenotypes (Figure 5.6B).

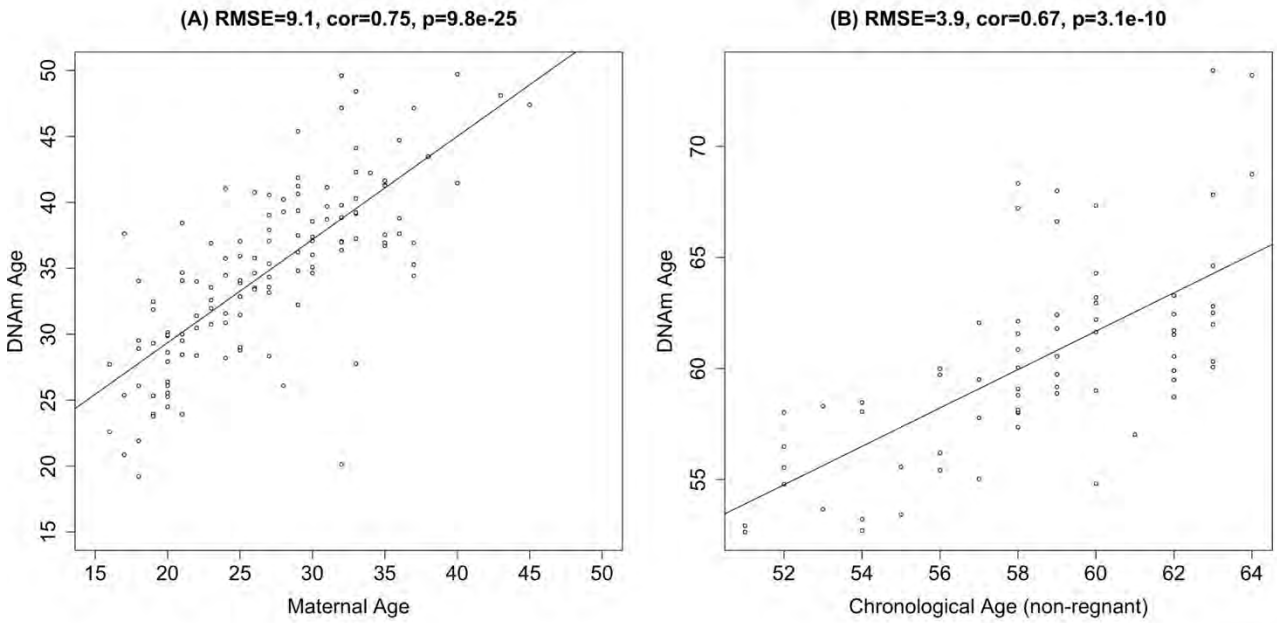

Figure 5.6 Pregnant women showed elevated methylation age compare to non-pregnant women. (A) DNA methylation age estimated for pregnant women. (B) DNA methylation age estimated for non-pregnant women.

5.4 Discussion

We describe for the first time, DNA methylation changes in maternal blood across 6-23 weeks' pregnancy. The study demonstrated that the DNA methylation profile did not change dramatically across early pregnancy. The subtle DNA methylation changes across early to mid-pregnancy were related to gestational age, smoking and maternal age. The cell proportions of blood samples from pregnant and non-pregnant women were different which may reflect immune tolerance that is essential for pregnancy success. In addition, the estimated DNA methylation age of pregnant women is higher than non-pregnant women in this study with the limitation that the publicly available non-pregnant women cohort is much older and likely to be post-menopausal.

The 5-cytosine methylated DNA is more stable than RNA and more sensitive to environmental stimuli than DNA itself [28, 29]. DNA methylation changes in maternal peripheral leukocytes could be promising biomarkers for pregnancy health. Using a rat model or human samples, previous studies have shown that DNA methylation changes in maternal blood could predict intrauterine growth retardation (IUGR) [30] and gestational diabetes mellitus (GDM) [31]. Our study showed that DNA methylation of maternal leukocytes slightly increased across early gestation which could be due to changes in the maternal immune system related to important physiological alterations and adaptations for maintenance of pregnancy [24].

Our study has confirmed data from others that smoking during pregnancy is associated with DNA methylation changes in maternal peripheral leukocytes [32]. Decreased DNA methylation (3 sites) was observed in *AHRR* gene body regions between women who smoked throughout pregnancy and non-smokers in DNA from the fetal side of the placenta [33]. Similarly, in maternal circulating leukocytes from our study, DNA methylation of one site we identified at *AHRR* gene body is significantly lower in women who were smoking at sampling compared to that in women who did not smoke. In a previous study, the *TSHZ3* gene also contains a CpG site with greater methylation in leukocytes of pregnant women who smoked [34]. We also found greater DNA methylation at that site in the *TSHZ3* gene, but this was not statistically significant.

GLIPR2 is hypomethylated in cord blood between newborns with and without exposure to maternal smoking *in utero* [35]. However, the reduced DNA methylation of *GLIPR2* was not significant in our study. *GPR1*-AS1 tended to have a different methylation pattern between smoking and non-smoking groups. *SLIT3* is a gene that functions during cell migration. The increased DNA methylation of a site in *SLIT3* could be a new indicator for maternal smoking according to our study.

Very few studies have revealed that DNA methylation changes in maternal blood were associated with gestational age. We have found that there is an overall subtle increase of DNA methylation across early to mid-pregnancy in the maternal blood profile. Our study showed that a DNA methylation site in the 5'UTR of the maternally expressed imprinted gene *GRB10* is negatively correlated with gestational age. Methylation of *GRB10* may be an indicator for fetal development throughout early gestation because this gene is also important for placenta development [36]. Another study also showed that increased DNA methylation changes of the *GRB10* gene in chorionic villous samples is associated with spontaneous abortion [37].

A previous study showed that early pregnancy may be characterized by widespread hypomethylation in leukocytes compared with non-pregnant states [12]. However, we did not observe overall hypomethylation of blood from pregnant women compared to non-pregnant women. This may be due to the major limitation of this study that is

maternal leukocyte DNA data from non-pregnant women were obtained from the GEO database, and the non-pregnant women from whom samples were derived were significantly older than the pregnant women in our study. The cell proportion analyses showed a decrease in CD4 T cells in blood from pregnant women compared to non-pregnant women. This has been shown in a previous study that the CD4 T cell population significantly decreased with advancing gestational age, indicating CD4 T cells play an important role in the maintenance of pregnancy [38, 39]. Our data also show that there is an increase in neutrophils in blood from pregnant women across early to mid-gestation compared to non-pregnant women, which is in concordance with the result from a previous study that the number of neutrophils increases and then plateaus in maternal blood across early to mid-pregnancy [40].

DNA methylation of *ASPA* and *TRIM59* has previously been shown to strongly correlate with age in blood samples [41, 42]. In this study, the detected sites related to maternal age showed decreased DNA methylation at TSS1500 of *ASPA* and an obvious increase of DNA methylation at TSS1500 of *TRIM59* gene These may indicate greater molecular aging than would be expected for the actual maternal age. A previous study has also shown that parity is associated with an increase in epigenetic age [43]. In the current study, we confirmed that DNA methylation in maternal leukocytes can be used to estimate age and we found that pregnant women had an elevated methylation age compared to non-pregnant women. This is particularly

interesting because the publicly available data for the non-pregnant women was derived from older likely post-menopausal women. Potentially, the molecular aging seen in pregnant women may be reversible after delivery. More aging markers and mechanisms of aging during pregnancy could potentially be identified in the future.

In summary, we found that DNA methylation of maternal leukocytes was associated with maternal age, maternal smoking and fetal gestational age. There are DNA methylation and methylation age differences between maternal leukocyte samples from pregnant and non-pregnant women. The cell proportions of maternal leukocyte samples from pregnant and non-pregnant women are also different. These phenomena could be associated with the adaptive immune response during pregnancy which is characterised by tolerance to paternal antigen [44].

5.5 Conclusion

In conclusion, this study provides evidence for DNA methylation changes in maternal circulating leukocytes across 6-23 weeks' pregnancy and in a large group of pregnant women. These changes in DNA methylation in early pregnancy is not dramatic, but they are associated with gestational age, maternal smoking and maternal age. We show changes in cell proportions and methylation age differences between pregnant and non-pregnant women which deserve further investigation and confirmation. This

study may provide insights about more biomarkers for aging and maternal smoking during early pregnancy.

5.6 Supplementary figures and tables

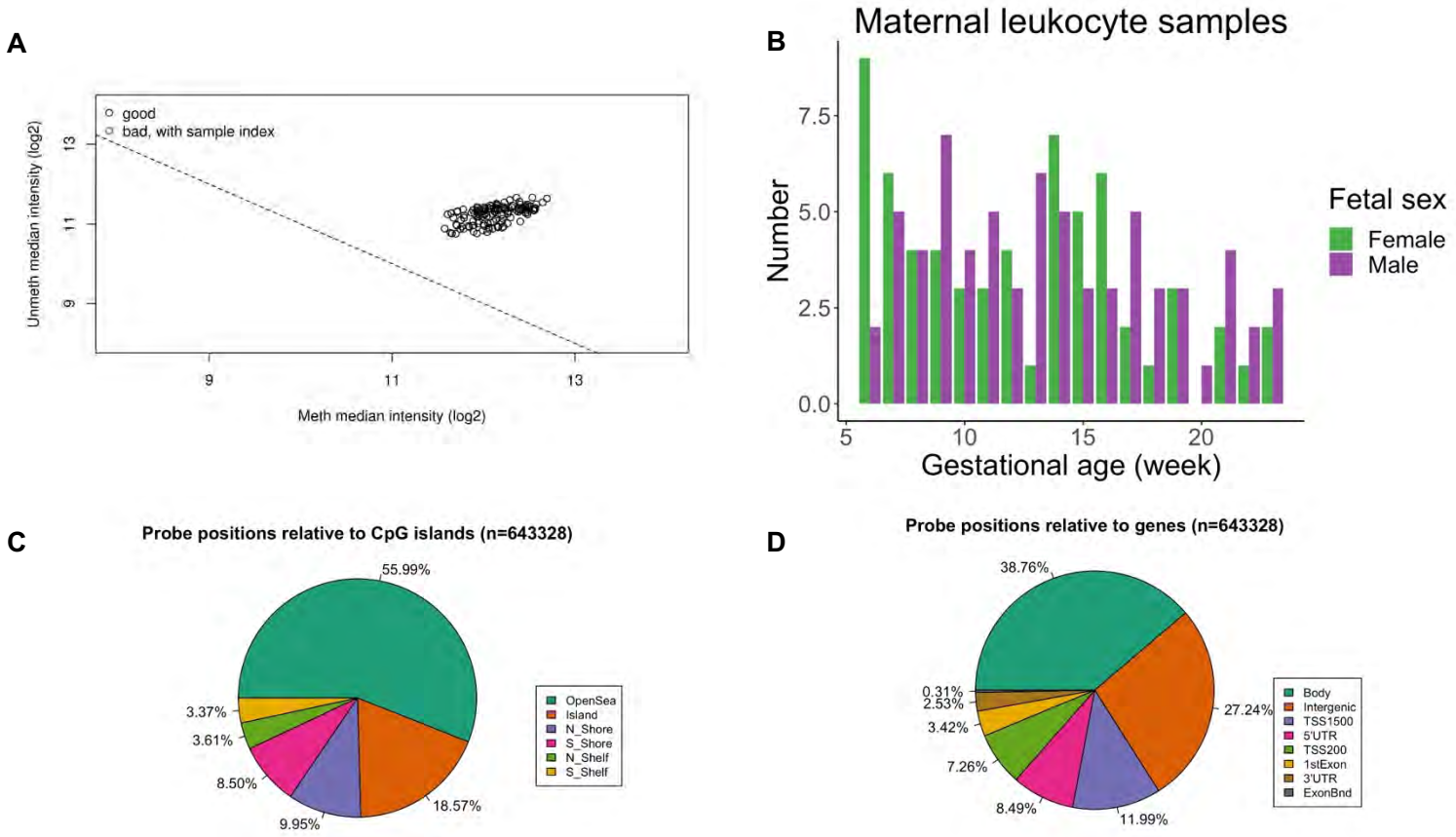

Figure S5.1 Quality control of maternal leukocyte samples and information of filtered probes. (A) Quality control plot showed that are maternal leukocyte samples (n=131) are of good quality. (B) Number of maternal leukocyte samples across early pregnancy with

fetal sex labelled. (C) Percentage of probes with position relative to CpG islands. (D) Percentage of probes with position relative to genes.

Table S.5.1 Meta data of maternal leukocyte samples (n=131).

Sample.ID	Array date received	Gestational age	Trimester	Fetal sex	Maternal age	Smoking status	Tissue
PAC0006	2017-03-23	9	First	F	32	N	Buffy coat
PAC0007	2017-03-23	10	First	M	28	N	Buffy coat
PAC0008	2017-03-23	9	First	F	26	N	Buffy coat
PAC0009	2017-03-23	8	First	F	30	Y	Buffy coat
PAC0010	2017-03-23	8	First	F	31	N	Buffy coat
PAC0011	2017-03-23	11	First	M	29	N	Buffy coat
PAC0012	2017-03-23	7	First	F	29	N	Buffy coat
PAC0013	2017-03-23	10	First	F	28	N	Buffy coat
PAC0014	2017-03-23	9	First	M	21	N	Buffy coat
PAC0015	2017-03-23	11	First	F	35	Y	Buffy coat
PAC0016	2017-03-23	9	First	M	27	Y	Buffy coat
PAC0017	2017-03-23	14	Second	F	27	N	Buffy coat
PAC0018	2017-08-01	21	Second	M	18	N	Buffy coat
PAC0020	2017-08-01	20	Second	M	27	Y	Buffy coat
PAC0021	2017-08-01	23	Second	M	45	N	Buffy coat
PAC0022	2017-08-01	19	Second	M	24	Y	Buffy coat
PAC0023	2017-08-01	22	Second	M	35	Y	Buffy coat
PAC0024	2017-03-23	6	First	F	31	Y	Buffy coat
PAC0025	2017-03-23	7	First	M	25	N	Buffy coat
PAC0026	2017-03-23	16	Second	M	18	N	Buffy coat
PAC0027	2017-03-23	16	Second	F	29	N	Buffy coat
PAC0029	2017-08-01	17	Second	M	19	N	Buffy coat
PAC0030	2017-03-23	13	First	M	32	Y	Buffy coat

PAC0031	2017-03-23	16	Second	M	20	N	Buffy coat
PAC0032	2017-03-23	13	First	M	21	N	Buffy coat
PAC0033	2017-08-01	22	Second	F	16	N	Buffy coat
PAC0034	2017-03-23	8	First	F	23	Y	Buffy coat
PAC0035	2017-03-23	6	First	F	27	N	Buffy coat
PAC0036	2017-03-23	6	First	F	23	N	Buffy coat
PAC0037	2017-03-23	7	First	F	22	N	Buffy coat
PAC0038	2017-03-23	9	First	F	29	N	Buffy coat
PAC0039	2017-03-23	6	First	F	19	N	Buffy coat
PAC0040	2017-03-23	13	First	M	27	Y	Buffy coat
PAC0041	2017-03-23	6	First	F	30	N	Buffy coat
PAC0042	2017-03-23	8	First	M	25	Y	Buffy coat
PAC0043	2017-03-23	12	First	M	19	Y	Buffy coat
PAC0044	2017-03-23	9	First	M	19	Y	Buffy coat
PAC0045	2017-03-23	7	First	M	24	N	Buffy coat
PAC0046	2017-03-23	8	First	F	43	N	Buffy coat
PAC0047	2017-03-23	10	First	M	22	N	Buffy coat
PAC0048	2017-03-23	6	First	M	33	N	Buffy coat
PAC0049	2017-03-23	11	First	M	27	N	Buffy coat
PAC0050	2017-03-23	9	First	M	20	N	Buffy coat
PAC0051	2017-03-23	7	First	F	26	N	Buffy coat
PAC0052	2017-03-23	9	First	M	33	N	Buffy coat
PAC0053	2017-03-23	16	Second	F	23	N	Buffy coat
PAC0054	2017-03-23	14	Second	M	29	N	Buffy coat
PAC0055	2017-03-23	6	First	F	20	N	Buffy coat
PAC0056	2017-03-23	6	First	F	21	Y	Buffy coat
PAC0057	2017-08-01	17	Second	M	26	N	Buffy coat
PAC0058	2017-03-23	16	Second	F	17	Y	Buffy coat

PAC0059	2017-03-23	10	First	F	25	Y	Buffy coat
PAC0060	2017-03-23	9	First	M	31	Y	Buffy coat
PAC0062	2017-03-23	10	First	M	29	Y	Buffy coat
PAC0063	2017-03-23	9	First	M	32	N	Buffy coat
PAC0064	2017-03-23	7	First	F	33	N	Buffy coat
PAC0065	2017-08-01	12	First	M	21	Y	Buffy coat
PAC0069	2017-08-01	15	Second	M	32	Y	Buffy coat
PAC0070	2017-08-01	13	First	M	27	Y	Buffy coat
PAC0071	2017-08-01	12	First	F	35	N	Buffy coat
PAC0072	2017-08-01	12	First	M	33	N	Buffy coat
PAC0074	2017-08-01	13	First	F	18	Y	Buffy coat
PAC0075	2017-08-01	23	Second	M	25	N	Buffy coat
PAC0076	2017-08-01	18	Second	M	21	N	Buffy coat
PAC0077	2017-08-01	19	Second	M	24	Y	Buffy coat
PAC0078	2017-08-01	14	Second	F	35	N	Buffy coat
PAC0083	2017-08-01	8	First	M	35	Y	Buffy coat
PAC0084	2017-08-01	19	Second	F	18	Y	Buffy coat
PAC0086	2017-08-01	11	First	M	20	N	Buffy coat
PAC0087	2017-08-01	14	Second	M	26	Y	Buffy coat
PAC0088	2017-08-01	11	First	F	22	N	Buffy coat
PAC0091	2017-08-01	11	First	M	23	Y	Buffy coat
PAC0093	2017-08-01	15	Second	M	30	Y	Buffy coat
PAC0097	2017-08-01	16	Second	F	NA	NA	Buffy coat
PAC0098	2017-08-01	14	Second	M	36	N	Buffy coat
PAC0099	2017-08-01	14	Second	F	26	Y	Buffy coat
PAC0100	2017-08-01	15	Second	F	40	Y	Buffy coat
PAC0102	2017-08-01	15	Second	F	25	Y	Buffy coat
PAC0103	2017-08-01	15	Second	M	40	N	Buffy coat

PAC0105	2017-08-01	12	First	F	19	N	Buffy coat
PAC0107	2017-08-01	10	First	F	25	Y	Buffy coat
PAC0108	2017-08-01	18	Second	M	20	Y	Buffy coat
PAC0109	2017-08-01	14	Second	F	17	N	Buffy coat
PAC0111	2017-08-01	16	Second	F	29	N	Buffy coat
PAC0114	2017-08-01	21	Second	M	36	Y	Buffy coat
PAC0117	2017-08-01	21	Second	M	32	N	Buffy coat
PAC0118	2017-08-01	23	Second	F	29	N	Buffy coat
PAC0120	2017-08-01	14	Second	M	22	N	Buffy coat
PAC0122	2017-08-01	10	First	M	27	N	Buffy coat
PAC0124	2017-08-01	17	Second	F	24	Y	Buffy coat
PAC0127	2017-08-01	18	Second	M	21	N	Buffy coat
PAC0129	2017-08-01	14	Second	M	37	Y	Buffy coat
PAC0131	2017-08-01	18	Second	F	24	Y	Buffy coat
PAC0134	2017-08-01	16	Second	F	17	N	Buffy coat
PAC0139	2017-08-01	22	Second	M	32	N	Buffy coat
PAC0140	2017-08-01	21	Second	F	28	Y	Buffy coat
PAC0019	2019-02-01	17	Second	M	33	N	Buffy coat
PAC0068	2019-02-01	13	First	M	35	N	Buffy coat
PAC0092	2019-02-01	13	First	M	30	Y	Buffy coat
PAC0094	2019-02-01	11	First	M	36	Y	Buffy coat
PAC0101	2019-02-01	14	Second	F	20	Y	Buffy coat
PAC0106	2019-02-01	11	First	F	20	N	Buffy coat
PAC0110	2019-02-01	15	Second	F	37	N	Buffy coat
PAC0113	2019-02-01	12	First	F	40	Y	Buffy coat
PAC0121	2019-02-01	14	Second	F	37	N	Buffy coat
PAC0125	2019-02-01	15	Second	F	21	Y	Buffy coat
PAC0126	2019-02-01	12	First	F	23	Y	Buffy coat

PAC0128	2019-02-01	14	Second	F	20	Y	Buffy coat
PAC0130	2019-02-01	17	Second	M	27	N	Buffy coat
PAC0135	2019-02-01	16	Second	M	29	Y	Buffy coat
PAC0137	2019-02-01	17	Second	M	18	N	Buffy coat
PAC0141	2019-02-01	8	First	M	26	Y	Buffy coat
PAC0148	2019-02-01	8	First	M	32	N	Buffy coat
PAC0183	2019-02-01	15	Second	F	16	N	Buffy coat
PAC0186	2019-02-01	7	First	M	32	N	Buffy coat
PAC0191	2019-02-01	19	Second	F	24	Y	Buffy coat
PAC0192	2019-02-01	21	Second	M	34	Y	Buffy coat
PAC0193	2019-02-01	6	First	F	30	N	Buffy coat
PAC0196	2019-02-01	21	Second	F	20	N	Buffy coat
PAC0198	2019-02-01	7	First	M	33	N	Buffy coat
PAC0200	2019-02-01	7	First	F	38	Y	Buffy coat
PAC0202	2019-02-01	6	First	M	25	Y	Buffy coat
PAC0203	2019-02-01	19	Second	F	18	Y	Buffy coat
PAC0204	2019-02-01	7	First	M	33	N	Buffy coat
PAC0208	2019-02-01	7	First	F	30	N	Buffy coat
PAC0211	2019-02-01	17	Second	F	20	Y	Buffy coat
PAC0214	2019-02-01	9	First	F	25	N	Buffy coat
PAC0215	2019-02-01	23	Second	M	37	N	Buffy coat
PAC0219	2019-02-01	6	First	F	19	N	Buffy coat
PAC0221	2019-02-01	19	Second	M	33	N	Buffy coat
PAC0222	2019-02-01	23	Second	F	25	N	Buffy coat

Table S.5.2 DNA methylation sites associated with PC1 and smoking.

Chromosome	Position	Name	UCSC RefGene name	UCSC RefGene group	DNA methylation (Smoker)	DNA methylation (Non-Smoker)	Deta Beta (smoker vs non-smoker)	Correlation with PC1 (correlation P value < 0.05)
chr5	373378	cg05575921	AHRR	Body	0.69474764	0.84895029	-0.1542026	-
chr5	168217234	cg19589725	SLIT3;SLIT3	Body;Body	0.41791186	0.316078	0.10183385	-0.3952281
chr14	101069717	cg10829391			0.74512331	0.63790793	0.10721538	-
chr7	30219135	cg06405219			0.38314947	0.28187733	0.10127214	-
chr19	31799406	cg18843803	TSHZ3	Body	0.87516121	0.77440518	0.10075603	-
chr12	58570196	cg24418762			0.59303987	0.49047384	0.10256603	-
chr5	29568333	cg00086247			0.250736	0.3532745	-0.1025385	-0.2188674
chr2	20871002	cg10687131	GDF7	Body	0.40862553	0.30855077	0.10007476	-
chr1	82627662	cg07748963			0.0959137	0.19745061	-0.1015369	-
chr3	42387524	cg10123377			0.50606187	0.38815799	0.11790388	-
chr22	30308934	cg26167301	MTMR3;MTMR3;MTMR3	5'UTR;5'UTR;5'UTR	0.538816936	0.429408057	0.109408878	-
chr10	31040939	cg20673407			0.45863502	0.56058259	-0.1019476	-
chr10	28781637	cg11671940			0.49292503	0.38225754	0.11066749	-
chr2	207090465	cg22444562	GPR1-AS	Body	0.61100422	0.46397567	0.14702855	-
chr15	85177792	cg15009596	SCAND2P;SCAND2P	Body;Body	0.70989044	0.81079845	-0.100908	-
chr9	36154750	cg19097407	GLIPR2;GLIPR2;GLIPR2;GLIPR2;GLIPR2; GLIPR2;GLIPR2;GLIPR2;GLIPR2;GLIPR2;GLIPR2	Body;Body;Body;Body;Body; Body;Body;Body;Body;Body;Body	0.29053625	0.393791373	-0.103255124	-0.327650482
chr7	76221076	cg17404449	LOC100133091	Body	0.68914548	0.54238839	0.14675709	-

Table S5.3 DNA methylation sites statistically significantly associated with PC1 and gestational age (GA).

Chromosome	Position	Name	UCSC RefGene name	UCSC RefGene group	SD of beta values across GA	Correlation with PC1	Correlation with GA	FDR (Correlation with GA)
chr7	50791421	cg12999836	GRB10;GRB10;GRB10	Body;Body;5'UTR	0.03603497	-0.3693107	-0.5079993	3.78E-04
chr7	27207318	cg22260008	HOXA10-AS;HOXA10-HOXA9	TSS1500;Body	0.03011008	-0.3668673	-0.473821	0.0070085
chr12	122442863	cg00052692			0.05167106	-0.3867188	-0.4513361	0.04044806
chr15	38784291	cg07232609	RASGRP1;RASGRP1;RASGRP1	Body;Body;Body	0.03282354	0.20762558	0.45500342	0.030651
chr15	101137253	cg25780496	LINS	5'UTR	0.01485858	-0.4148818	-0.4679995	0.01116902
chr3	101901234	cg12992827			0.05769898	-0.3097536	-0.4924339	0.00148793
chr16	11299766	cg08876198			0.03619155	-0.3957219	-0.4543004	0.03233287
chr8	23405208	cg14505131	SLC25A37	Body	0.03628865	-0.207491	-0.4624741	0.01724443
chr1	1689452	cg04439557	NADK	Body	0.01743698	-0.382237	-0.465338	0.01378148
chr19	17959783	cg20959703	JAK3	TSS1500	0.03558119	-0.3474714	-0.4530578	0.03552421
chr19	4973177	cg15745450	KDM4B	5'UTR	0.03440658	-0.19799	-0.4596803	0.02141736
chr16	60804969	cg05263979			0.03469328	-0.296304	-0.4606094	0.01993242
chr3	185538892	cg24960291	IGF2BP2;IGF2BP2	Body;Body	0.02920448	-0.3027254	-0.4629116	0.01666619

Table S5.4 DNA methylation sites statistically significantly associated with PC3 and maternal age (MA).

Chromosome	Position	Name	UCSC RefGene name	UCSC RefGene group	SD of beta values across MA	Correlation with PC3	Correlation with MA	FDR (Correlation with MA)
chr1	68512845	cg13697378	DIRAS3	Body	0.036652	0.26563765	0.52298274	1.12E-04
chr17	28562685	cg14692377	SLC6A4;SLC6A4	1stExon;5'UTR	0.02025666	0.24130829	0.48064621	0.00459227
chr14	93389628	cg21620282	CHGA;CHGA;CHGA;CHGA	5'UTR;5'UTR;1stExon;1stExon	0.02871193	-0.1742736	0.48641464	0.00284966
chr6	35490818	cg18468088			0.0224442	0.21868405	0.45317444	0.03963178
chr3	160167977	cg07553761	TRIM59	TSS1500	0.05190714	0.23454506	0.49798456	0.00106509
chr1	40098811	cg12634306	HEYL	Body	0.05356098	0.21071845	0.48105268	0.00444173
chr18	66389420	cg19283806	CCDC102B	5'UTR	0.05068091	0.18743619	-0.4830347	0.003773
chrX	53741945	cg08975875			0.03668165	0.26692342	-0.4517365	0.04413713

chr3	137838021	cg20747538			0.02316721	0.2992941	-0.4572363	0.02915982
chr12	7245510	cg23368715	C1R	TSS1500	0.02328596	0.233511	-0.4557206	0.0327124
chr2	62683973	cg14887028			0.01863478	0.18158736	-0.4914856	0.00185961
chr3	14832729	cg07504615			0.02384293	0.19143198	-0.6209567	2.09E-09
chrX	49166019	cg15833111	GAGE10	Body	0.03363612	0.28918662	-0.4809691	0.00447228
chr20	4721316	cg27037708	PRNT;PRNT;PRNT	TSS200;TSS200;TSS200	0.02399583	0.24600745	-0.5083836	4.26E-04
chr15	101971587	cg18767232	PCSK6;PCSK6;PCSK6;PCSK6;PCSK6;PCSK6;PCSK6	Body;Body;Body;Body;Body;Body;Body	0.04245102	0.27197619	-0.4603743	0.02294261
chr17	48193712	cg24087280	SAMD14;SAMD14	Body;Body	0.03094687	0.21965982	-0.4531332	0.03975458
chr11	35638398	cg11741201	FJX1	TSS1500	0.02569574	0.24275803	-0.4664398	0.01433345
chr1	241970344	cg17646797			0.01571006	0.21420654	-0.4518561	0.04374469
chr7	140739534	cg12776916			0.03803345	0.17812132	-0.5543547	5.01E-06
chr1	68512807	cg20149168	DIRAS3	Body	0.0491724	0.24592111	0.5075193	4.60E-04
chr7	24705461	cg07050404	MPP6;MPP6	Body;Body	0.0251897	0.19677646	-0.5061341	5.21E-04
chr13	51723181	cg24674215			0.0254417	0.27278252	-0.468057	0.01262426
chr16	69743438	cg09286894	NQO1;NQO1;NQO1;NQO1	3'UTR;3'UTR;3'UTR;3'UTR	0.02807307	-0.2586461	-0.4935537	0.00155936
chr1	203474297	cg11122885	OPTC	3'UTR	0.02187406	0.20881757	-0.4876368	0.00257273
chr11	86517110	cg00475490	PRSS23;PRSS23;PRSS23;PRSS23;PRSS23;PRSS23;PRSS23	5'UTR;Body;Body;5'UTR;Body;Body;Body	0.03670527	0.17935465	-0.464369	0.01684793
chr1	241774637	cg19501755	OPN3	Body	0.03307735	0.18538599	-0.4974158	0.00111885
chr1	213898865	cg18732257			0.02880666	0.2135547	-0.5085949	4.18E-04
chr7	19183280	cg17508941			0.0382908	0.22723839	0.46895504	0.01176152
chr17	48623772	cg20067719	SPATA20	TSS1500	0.03473553	0.34335426	-0.4760399	0.00668007
chr2	71098542	cg22112832			0.03656076	0.23700167	-0.5169876	1.95E-04
chr10	73287051	cg01029119	CDH23;CDH23;CDH23;CDH23;CDH23	Body;Body;Body;Body;Body	0.02108544	0.19822179	-0.4630294	0.0186942
chr9	116570319	cg14257429			0.02401317	0.18300535	-0.4693008	0.01144461
chr2	145116633	cg18826637			0.08312498	0.19254	-0.4955333	0.00131604
chr4	10583796	cg18805621	CLNK	Body	0.04162366	0.21578176	-0.4706051	0.01032152
chr3	168185313	cg07323488	EGFEM1P	Body	0.04916146	0.19594492	-0.5026689	7.08E-04

chr7	28218686	cg07522171	JAZF1-AS1;JAZF1	TSS1500;Body	0.02348807	0.23010515	-0.4644364	0.01675987
chr10	25463350	cg13206721	GPR158;GPR158-AS1	TSS1500;Body	0.02928301	0.20762818	0.453183	0.03960638
chr20	36850842	cg17015290	KIAA1755;KIAA1755	ExonBnd;Body	0.02201488	0.23317695	-0.4891737	0.00226105
chr9	139776929	cg14620941			0.03201029	0.25372572	-0.4679832	0.01269779
chr1	116563487	cg08737116	SLC22A15	Body	0.03336434	0.33591807	-0.4630626	0.01864612
chr16	58768477	cg10140957	GOT2	TSS1500	0.02463061	0.17771847	-0.5166614	2.01E-04
chr7	121944506	cg25045526	FEZF1;FEZF1;FEZF1;FEZF1;FEZF1-AS1	1stExon;1stExon;5'UTR;5'UTR;Body	0.02332009	0.19965215	0.45510737	0.0342644
chr6	140762703	cg26413501			0.03543517	-0.1978936	-0.4748152	0.00737306
chr1	87666926	cg00420212			0.04789914	0.19017523	-0.5012512	8.01E-04
chr17	3377797	cg18769729	ASPA;ASPA	5'UTR;TSS1500	0.03873181	-0.2590605	-0.5064468	5.06E-04
chr2	97202260	cg02085953	ARID5A	TSS1500	0.02325416	0.22035326	-0.513938	2.58E-04
chr2	64870490	cg10184289	SERTAD2	5'UTR	0.04737244	0.20141485	0.46950306	0.01126311
chrX	45616094	cg14817951			0.06193975	0.24060755	-0.4888492	0.00232367
chr14	105955879	cg02395812	C14orf80	TSS1500	0.01984611	0.18000643	-0.4615331	0.02098561
chr17	40177415	cg06874016	NKIRAS2;NKIRAS2;NKIRAS2;NKIRAS2;NKIRAS2	3'UTR;3'UTR;3'UTR;3'UTR;3'UTR	0.02694696	0.23779381	-0.5356641	3.31E-05
chr3	147126206	cg16181396	ZIC1	TSS1500	0.01847959	0.20054019	0.46310175	0.01858974
chr4	17387235	cg12318914			0.03925808	0.20829076	-0.547032	1.06E-05
chr1	46309093	cg27447795	MAST2	Body	0.0460074	0.19999656	-0.5657072	1.50E-06
chr1	155959036	cg24964298			0.02453034	0.23129906	-0.4655342	0.01538538
chr6	116883080	cg05389014			0.02255583	0.28231837	-0.5017489	7.67E-04
chr12	66089473	cg18215449			0.05336194	0.18660924	-0.5459545	1.19E-05
chr19	33182526	cg02283691	NUDT19	TSS1500	0.01365148	0.25902299	-0.4855438	0.00306425
chr3	52738881	cg27312987	GLT8D1;SPCS1;GLT8D1;GLT8D1;GLT8D1;GLT8D1	TSS1500;TSS1500;5'UTR;5'UTR;5'UTR;5'UTR	0.0229574	0.31971629	-0.4509803	0.04669924
chr9	136023407	cg03474926	RALGDS	Body	0.01746408	0.22966339	-0.5108794	3.40E-04

References

1. Locke WJ, Guanzon D, Ma C, Liew YJ, Duesing KR, Fung KY, et al. DNA methylation cancer biomarkers: Translation to the clinic. Frontiers in Genetics. 2019;10.
2. Nebbioso A, Tambaro FP, Dell'Aversana C, Altucci L. Cancer epigenetics: moving forward. PLoS genetics. 2018;14(6).
3. Ordovás JM, Smith CE. Epigenetics and cardiovascular disease. Nature Reviews Cardiology. 2010;7(9):510.
4. Luo C, Hajkova P, Ecker JR. Dynamic DNA methylation: In the right place at the right time. Science. 2018;361(6409):1336-40.
5. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature Reviews Genetics. 2012;13(7):484-92.
6. Bianco-Miotto T, Mayne BT, Buckberry S, Breen J, Lopez CMR, Roberts CT. Recent progress towards understanding the role of DNA methylation in human placental development. Reproduction (Cambridge, England). 2016;152(1):R23-R30.
7. Schuster J, Uzun A, Stablia J, Schorl C, Mori M, Padbury JF. Effect of prematurity on genome wide methylation in the placenta. BMC medical genetics. 2019;20(1):116.
8. Kazmi N, Sharp GC, Reese SE, Vehmeijer FO, Lahti J, Page CM, et al. Hypertensive Disorders of Pregnancy and DNA Methylation in Newborns: Findings From the Pregnancy and Childhood Epigenetics Consortium. Hypertension. 2019;74(2):375-83.

9. Enquobahrie DA, Moore A, Muhie S, Tadesse MG, Lin S, Williams MA. Early pregnancy maternal blood DNA methylation in repeat pregnancies and change in gestational diabetes mellitus status—a pilot study. Reproductive sciences. 2015;22(7):904-10.

10. Berdasco M, Esteller M. Clinical epigenetics: seizing opportunities for translation. Nature Reviews Genetics. 2019;20(2):109-27.

11. Shridhar V, Chu T, Simhan H, Shaw PA, Peters DG. High Resolution Analysis of the Human Placental Dna Methylome in Early Gestation. Prenatal Diagnosis. 2020.

12. White WM, Brost BC, Sun Z, Rose C, Craici I, Wagner SJ, et al. Normal early pregnancy: a transient state of epigenetic change favoring hypomethylation. Epigenetics. 2012;7(7):729-34.

13. Miller S, Dykes D, Polesky H. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic acids research. 1988;16(3):1215.

14. Hanna CW, Peñaherrera MS, Saadeh H, Andrews S, McFadden DE, Kelsey G, et al. Pervasive polymorphic imprinted methylation in the human placenta. Genome research. 2016;26(6):756-67.

15. Zaimi I, Pei D, Koestler DC, Marsit CJ, De Vivo I, Tworoger SS, et al. Variation in DNA methylation of human blood over a 1-year period using the Illumina MethylationEPIC array. Epigenetics. 2018;13(10-11):1056-71.

16. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome biology. 2016;17(1):208.

17. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the

analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-9.

18. Horvath S. DNA methylation age of human tissues and cell types. Genome biology. 2013;14(10):3156.

19. Ritchie ME, Diyagama D, Neilson J, van Laar R, Dobrovic A, Holloway A, et al. Empirical array quality weights in the analysis of microarray data. BMC bioinformatics. 2006;7(1):261.

20. McGregor K, Bernatsky S, Colmegna I, Hudson M, Pastinen T, Labbe A, et al. An evaluation of methods correcting for cell-type heterogeneity in DNA methylation studies. Genome biology. 2016;17(1):84.

21. Pidsley R, Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC genomics. 2013;14(1):293.

22. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC bioinformatics. 2012;13(1):86.

23. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlén S-E, Greco D, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PloS one. 2012;7(7).

24. Gruzieva O, Merid SK, Chen S, Mukherjee N, Hedman AM, Almqvist C, et al. DnA Methylation Trajectories During Pregnancy. Epigenetics insights. 2019;12:2516865719867090.

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

26. Xu W, Xu M, Wang L, Zhou W, Xiang R, Shi Y, et al. Integrative analysis of DNA methylation and gene expression identified cervical cancer-specific diagnostic biomarkers. Signal Transduction and Targeted Therapy. 2019;4(1):1-11.

27. Titus AJ, Gallimore RM, Salas LA, Christensen BC. Cell-type deconvolution from DNA methylation: a review of recent applications. Human molecular genetics. 2017;26(R2):R216-R24.

28. Blair JD, Yuen RK, Lim BK, McFadden DE, von Dadelszen P, Robinson WP. Widespread DNA hypomethylation at gene enhancer regions in placentas associated with early-onset pre-eclampsia. Molecular human reproduction. 2013;19(10):697-708.

29. Ivorra C, Fraga MF, Bayón GF, Fernández AF, Garcia-Vicent C, Chaves FJ, et al. DNA methylation patterns in newborns exposed to tobacco in utero. Journal of translational medicine. 2015;13(1):25.

30. Wu D-M, Yan Y-E, Ma L-P, Liu H-X, Qu W, Ping J. Intrauterine growth retardation-associated syncytin b hypermethylation in maternal rat blood revealed by DNA methylation array analysis. Pediatric research. 2017;82(4):704-11.

31. Wu P, Farrell WE, Haworth KE, Emes RD, Kitchen MO, Glossop JR, et al. Maternal genome-wide DNA methylation profiling in gestational diabetes shows distinctive disease-associated changes relative to matched healthy pregnancies. Epigenetics. 2018;13(2):122-8.

32. Reese SE, Zhao S, Wu MC, Joubert BR, Parr CL, Håberg SE, et al. DNA methylation score as a biomarker in newborns for sustained maternal smoking during pregnancy. Environmental health perspectives. 2017;125(4):760-6.

33. van Otterdijk SD, Binder AM, Michels KB. Locus-specific DNA methylation in the placenta is associated with levels of pro-inflammatory proteins in cord blood and they are both independently affected by maternal smoking during pregnancy. Epigenetics. 2017;12(10):875-85.

34. Richardson TG, Richmond RC, North T-L, Hemani G, Davey Smith G, Sharp GC, et al. An integrative approach to detect epigenetic mechanisms that putatively mediate the influence of lifestyle exposures on disease susceptibility. International journal of epidemiology. 2019;48(3):887-98.

35. Howe CG, Zhou M, Wang X, Pittman GS, Thompson IJ, Campbell MR, et al. Associations between Maternal Tobacco Smoke Exposure and the Cord Blood CD 4+ DNA Methylome. Environmental health perspectives. 2019;127(4):047009.

36. Litzky JF, Deyssenroth MA, Everson TM, Armstrong DA, Lambertini L, Chen J, et al. Placental imprinting variation associated with assisted reproductive technologies and subfertility. Epigenetics. 2017;12(8):653-61.

37. Liu Y, Tang Y, Ye D, Ma W, Feng S, Li X, et al. Impact of abnormal DNA methylation of imprinted loci on human spontaneous abortion. Reproductive Sciences. 2018;25(1):131-9.

38. Seol H-J, Oh M-J, Lim J-E, Jung N-H, Yoon S-Y, Kim H-J. The role of CD4+ CD25bright regulatory T cells in the maintenance of pregnancy, premature rupture of membranes, and labor. Yonsei medical journal. 2008;49(3):366-71.

39. Santner-Nanan B, Peek MJ, Khanam R, Richarts L, Zhu E, de St Groth BF, et al. Systemic increase in the ratio between Foxp3+ and IL-17-producing CD4+ T cells in healthy pregnancy but not in preeclampsia. The Journal of Immunology. 2009;183(11):7023-30.

40. Bakrim S, Motiaa Y, Ouarour A, Masrar A. Hematological parameters of the blood count in a healthy population of pregnant women in the Northwest of Morocco (Tetouan-M'diq-Fnideq provinces). Pan African Medical Journal. 2018;29(1):1-12.

41. Bekaert B, Kamalandua A, Zapico SC, Van de Voorde W, Decorte R. Improved age determination of blood and teeth samples using a selected set of DNA methylation markers. Epigenetics. 2015;10(10):922-30.

42. Jung S-E, Lim SM, Hong SR, Lee EH, Shin K-J, Lee HY. DNA methylation of the ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59 genes for age prediction from blood, saliva, and buccal swab samples. Forensic Science International: Genetics. 2019;38:1-8.

43. Kresovich JK, Harmon QE, Xu Z, Nichols HB, Sandler DP, Taylor JA. Reproduction, DNA methylation and biological age. Human Reproduction. 2019;34(10):1965-73.

44. Dekker G, Robillard PY, Roberts C. The etiology of preeclampsia: the role of the father. Journal of reproductive immunology. 2011;89(2):126-32.

Statement of Authorship

Title of Paper	DNA methylation profile of maternal peripheral leukocytes across early to mid-pregnancy
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Qianhui Wan		
Contribution to the Paper	Performed the data analyses and wrote the manuscript.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	23/03/20

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	James Breen		
Contribution to the Paper	Conceived the study design, supervised the work, assisted in coding and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Shalem Yiner-Lee Leemaqz		
Contribution to the Paper	Assisted with statistics and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Tanja Jankovic-Karasoulos		
Contribution to the Paper	Laboratory work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Dale Christopher McAninch		
Contribution to the Paper	Laboratory work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Dylan McCullough		
Contribution to the Paper	Sample collection and laboratory work.		
Signature		Date	23/03/20

Name of Co-Author	Melanie Denise Smith		
Contribution to the Paper	Assisted in coding and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Tina Bianco-Miotto		
Contribution to the Paper	Conceived the study design, supervised the work and edited the manuscript.		
Signature		Date	23/03/20

Name of Co-Author	Claire Trelford Roberts		
Contribution to the Paper	Conceived the study design, supervised the work and edited the manuscript.		
Signature		Date	23/03/20

Please cut and paste additional co-author panels here as required.

6 General Discussion

The aim of this research was to profile DNA methylation of the placenta and matched maternal leukocytes across early gestation using robust bioinformatics methods. The novel findings of this study include developing a way to check tissue purity for placenta samples before downstream analyses. This is the first study to profile DNA methylation of placenta and matched maternal leukocytes from 6-23 weeks' gestation at a genome-wide scale and the first to report DNA methylation changes across the 10-11 weeks' gestation threshold when maternal blood presumably starts to flow into the placental intervillous space. Furthermore, we characterised DNA methylation changes in maternal blood across early gestation and found an association between fetal gestational age and the DNA methylation of maternal leukocytes in this study. Findings from this study contribute to bioinformatics in the field of human placental research, epigenetic profiling studies and biomarker identification.

6.1 Overall Significance

6.1.1 Appropriate bioinformatics methodologies and pipelines are important for data processing and interpretation

Currently, some studies use the default parameter of bioinformatics methods to detect differential methylation between sample groups. However, recent research [1] and our study, show that the results change dramatically when using different parameters for the same data set. This fact raises the issue that choosing the right methods and parameters are important for data interpretation from bioinformatics analyses. It is recommended to use a known or published data set to first test the overall

performance of different methods and then identify the most suitable method and parameters for the data set that the researchers wish to study as documented in Chapter 2.

Placenta has a different DNA methylation profile compared with other healthy human tissues, so a pipeline more suitable for analysing placenta data is required. Previous studies have documented some pipelines designed for DNA methylation array analyses, including the cross-package Bioconductor workflow [2] that demonstrated the analyses for 450K array data using several R packages (*minfi*, *limma*, *missMethyl*, *Gviz* and *DMRcate*). *RnBeads* is another comprehensive package that lately updated for analysing DNA methylation data from different platforms (450K array, EPIC array and WGBS) and it has a graphical user interface (GUI) named *RnBeadsDJ* [3]. *ChAMP* is also used for DNA methylation array analyses and have GUI for differential methylation analyses [4].

However, these pipelines mentioned above do not contain some analyses specifically for placenta such as analyses for PMDs and ICRs, the pipeline used in Chapter 4 will be more suitable for profiling DNA methylation in placenta. Using the pipeline developed during this study, researchers studying placental development can test DNA methylation changes in promoters, enhancers, PMDs and ICRs and do analyses such as detecting DMRs with proper parameters, predicting samples from mixed tissues that are not comprehensively included in previous pipelines, so this work filled the gap of placental data analyses using EPIC array data and contributed to choosing proper bioinformatic methods for analysing data from placental tissue. This pipeline

will be wrapped in a R package in future so that these methods can be used by other researchers conveniently.

6.1.2 Sample outliers need to be removed with caution

Outliers are commonly found in genome-wide DNA methylation studies. However, only some of the studies discussed why these outliers are dramatically different from other samples before removing them for data analyses [5]. PCA of our own studies and other tissues was used for identifying outliers and confirming whether these outliers are caused by the mixing of placenta tissue with other tissue types such as maternal decidua tissue. Using placenta specific features such as partially methylated domains (40-70% methylated) [6] and placenta-specific imprinting control regions (50% methylated) [7] we can identify placenta samples that are contaminated with other tissue types. Consequently, the outliers in our study can be removed before downstream analyses.

It is important to remove outliers with caution because there may be many factors such as phenotypes not included in meta data that cause the observed changes in the DNA methylation profile of these outliers. This is especially the case for placenta samples collected from early gestation when the outcomes of the fetuses and pregnancies are unknown. In Chapter 3, we documented the details about how to identify mixed placenta samples and a R shiny app will be developed and shortly available for other researchers use our method to check the features of outliers in their data.

6.1.3 DNA methylation changes in placenta across early pregnancy is associated with placental development

EPIC array data of 125 placenta samples from 6-23 weeks' gestation were used in this study to identify DNA methylation changes across early gestation. In Chapter 4, We first confirmed that DNA methylation increases gradually across early gestation. The hypomethylation at placental enhancers compared to non-placental enhancers are interesting and deserve further investigation, as there are studies indicating the important function of DNA methylation at enhancers during developmental processes [8, 9]. Genes differentially methylated between groups up to and after 10 weeks' gestation included *SLC2A3* and *CCR7* that were also differentially expressed, which could be involved in glucose transport to the fetus [10], immune tolerance [11] and differentiation of cytotrophoblasts [12] which are important processes during early pregnancy. In addition, the DNA methylation is stable at partially methylated domains and imprinting control regions across early gestation with very few DMRs in these regions. This confirms the establishment of partially methylated domains (PMDs) and gene imprinting as early as 6 weeks' gestation in placenta, which could be associated with the fundamental functions of placenta during early development.

This is one of the most comprehensive studies of the placental methylome across early gestation that could lay a foundation for further research on placental health and disease. First, compared to studies using 450K arrays, this study used EPIC array data with more DNA methylation sites detected [13]. Second, samples used in this study have gestational age covered week by week between 6-23 weeks' gestation. These data are valuable resources for comprehensive DNA methylation

characterisation across early human gestation. Third, the influence of oxygen on DNA methylation changes during early placental development was explored in this study. Previous studies demonstrated that oxygen is an important factor during placental development because cell differentiation and invasion may be regulated by oxygen gradients (i.e. placenta develops in an environment changing from low oxygen to normal oxygen) [14-16]. In this study, DMP, DMR and DMB between the two groups separated according to the oxygen tension changes (approximately 10 weeks') across early gestation were identified and the genes related with these regions were investigated. This adds to the understanding of the methylome of placenta and highlights the need for future research to verify and uncover the functions of these DNA methylation changes.

6.1.4 Maternal leukocytes are sources of biomarkers for both maternal and fetal health status

As placenta samples are hard to obtain during early gestation and chorionic villous sampling for clinical tests may cause miscarriage [17], a good way to monitor pregnancy health could be using maternal blood samples which is generally considered non-invasive. In Chapter 5, we investigated the DNA methylation of the maternal leukocyte samples from the same fetal/maternal pairs where placenta samples were collected. Unlike DNA methylation of placenta, DNA methylation of maternal leukocytes did not show a gradual changing pattern across early pregnancy, but some site-specific changes were detected that were associated with maternal smoking, maternal age and gestational age.

The study of DNA methylome of maternal leukocytes here contributes to the field of biomarker identification during early pregnancy. Frist, unlike methylation of cell-free DNA derived from the fetus [18], using matched maternal leukocytes, no DMP nor DMRs were identified between first and second trimester groups in this study. This is a normal phenomenon since the DNA methylation changes remains probe-wise with a low number of probes even for control and disease comparison groups in maternal blood [19, 20]. In future, single cell techniques will be suitable for investigating the DNA methylation changes in different maternal leukocyte populations across gestation and have potential for biomarker identification. Second, this is the first study that undertook correlation analyses between DNA methylation of each site and fetal gestational age. We reported that changes of DNA methylation at 13 sites in maternal leukocytes were statistically significantly correlated with gestational age. These sites will be further verified by pyrosequencing [21] in future to confirm whether they are real biomarkers for normal placental development.

6.2 Limitations of this study

6.2.1 Limitation of EPIC array data

DNA methylation arrays are a more targeted and cost-effective platform for detecting DNA methylation of the human genome. Compared to 450K array, the number of DNA methylation sites is nearly doubled, especially in open sea regions of the genome in EPIC array, however, the coverage of the EPIC array is still lower compared with WGBS [22]. Although EPIC array contains more than 800,000 probes and data obtained from EPIC array are informative enough to show the patterns of DNA methylation in placenta and maternal peripheral leukocytes, it may not contain probes

in some regions that may be important for early placental development, so replication of these results from EPIC arrays using WGBS data in future will be important to do. WGBS is still the gold standard for assessing whole genome DNA methylation profiles. WGBS assessment will provide a more rigorous means to verify the findings in this study in future, which will also increase the chances of identifying more site-based biomarkers in profiling the DNA methylation of maternal leukocytes.

6.2.2 Limitation of R coding for big data analyses

The R programming language is an open source language and contains a lot of packages suitable for genome and epigenome studies. The main limitation of R programming language is that R is slow in processing large matrices or large data sets, especially for calculating correlations for large matrices or computing loops since it uses too much computer random-access memory (RAM) [23]. R packages and R programming are much slower than other languages such as C programs [24, 25]. We used R for data analyses in this study because EPIC array data is relatively smaller in size compared with WGBS data. Programmers are always trying their best to improve the computational efficiency and reduce computational load using different strategies (e.g. *Rcpp* R package that allows the implementation of C codes in R scripts) while using R. In future, splitting of input data or applying reasonable filters to obtain a smaller input dataset may be helpful to increase the speed analyses in R. Furthermore, other programming languages such as C and Python will also be used in future for analyses with larger data sets such as those generated by WGBS.

6.2.3 Limitation of non-pregnant cohort for maternal leukocyte DNA methylation study

Since there is not any published DNA methylation data (EPIC array) for leukocyte samples from non-pregnant women matching the age of the pregnant women in our cohort, non-age matched data from GEO data base are used as documented in Chapter 5. The differences in DNA methylation between leukocytes from pregnant and non-pregnant women in our study may be influenced by the different age ranges of the two groups of women because DNA methylation is increased in older adults compared to young adults [26]. The public control dataset of circulating leukocytes from younger non-pregnant women that we found is a dataset from the 27K platform [27]. The number of probes for this dataset is too low compared to EPIC array data, so it was not useful for this study. DNA methylation of leukocyte samples from women of childbearing age and matched with the age range of the pregnant women in our cohort will be profiled in future to verify the results in Chapter 5.

6.3 Future directions

As human placenta is the least understood human organ and is uniquely hypomethylated compared with other human tissues, there are still a lot of unknown aspects that need to be investigated. This study detailed data analysis methods and explored DNA methylation profiles of placenta and matched maternal leukocytes in early pregnancy to study placental development and pregnancy health. Except for the work presented in this study, there are many more important features of placental and maternal DNA methylation that can potentially be explored based on this work.

First, fetal sex differences between placenta samples from male and female bearing pregnancies can be of great importance because huge differences were observed between DNA methylation from placenta in pregnancies carrying male versus female fetuses, especially on the sex chromosomes (data not included in this thesis). Sex chromosomes were removed when profiling DNA methylation of placenta samples in the current study (Chapter 4) to minimise the influence of sex differences. Consequently, we are yet to investigate sex differences in the placental DNA methylation profile in early pregnancy. In future, analyses with sex chromosomes included will be performed and RNA sequencing and microRNA sequencing data will be integrated to identify sex differences between placenta from male and female bearing pregnancies.

As this research is part of a large NIH funded project that is characterising the methylome, transcriptome and miRnome of the human placenta across gestation, the integration of DNA methylation, RNA sequencing and microRNA sequencing data from placenta and maternal blood will be performed in the near future with up-to-date methodologies potentially including machine learning algorithms to characterise early placental development more comprehensively. This could provide details about fetal sex differences across “omics” during early gestation.

The unique placenta DNA methylation profile indicates more appropriate analytic methods and pipelines, integrated analyses and better data visualisation and interpretation are required for this essential developmental tissue. According to Chapter 2, to better optimise models’ parameters before using our own data, in future,

we can adopt machine learning approaches and split data into training and test datasets. For future studies, we also recommend always considering fetal sex differences [28, 29] and cell compositions [30, 31] in the analytic pipelines for data from placenta and maternal leukocytes. Following Chapter 4 and Chapter 5, in future, we aim to identify more associations between DNA methylation changes across gestation and the functions of these changes during pregnancy. This could be done using WGCNA analyses (see below) and examining DNA methylation by gene-set analysis. If these changes are important for human pregnancy, whether they can be reflected by DNA methylation changes in maternal blood, which can benefit the non-invasive monitoring of pregnancy [32]. A recent study has used machine learning methods to predict DNA methylation levels in placenta using the DNA methylation profile of cord blood [33], which indicates that machine learning methods could be useful for data integration and biomarker identification. This of course provides information just at the time of birth. Furthermore, better data integration and visualisation will also contribute to better data interpretation in future work [34, 35].

When plotting the DNA methylation changes of some differentially methylated regions across early gestation, small and gradual changes were observed. However, these patterns were not well captured by comparing the DNA methylation differences between two sample groups. Better ways to find small and consistent changes of DNA methylation across samples or across gestation will be required. Weighted correlation network analysis (WGCNA) proved to be a good way to identify DNA methylation patterns associated with a continuous variable [36]. Actually, we have tested the WGCNA method for co-methylation analyses using the placenta data in this study (not included in this thesis) but different from 450K arrays [37], the correlation calculation

step for the large numbers of probes on EPIC arrays is too computationally intensive in R because of the large number of probes in the EPIC array. In future, an optimised filtering of probes (likely by filtering out more unchanged probes across samples) needs to be applied before WGCNA analyses to investigate the co-methylation patterns in placenta across early gestation. Through co-methylation analyses, DNA methylation of gene promoters or other regulatory elements that are co-methylated across early gestation will be identified, which may help better elucidate the time-series changing of DNA methylation during early placental development.

The epigenetic clock for different types of human tissue is an important tool to estimate their DNA methylation age. In Chapter 4, we established an epigenetic clock for predicting gestational age with placenta DNA methylation data and by using the existing Horvath clock [38], and predicted maternal DNA methylation age. This study showed that DNA methylation changes in maternal leukocytes associated with gestational age. A recent study also demonstrated gestational age prediction (placenta epigenetic clock) using public DNA methylation array data from a large number of placenta samples (1102 as training data and 187 as test data) [39]. In future, an epigenetic clock that uses DNA methylation of maternal leukocytes to estimate gestational age should be explored. This epigenetic clock could be a useful tool to estimate the missing meta data of publicly available data and may be useful in studies of pregnancy complications.

Placenta is a heterogeneous tissue that contains numerous cell types. Ideally, future work could consider profiling DNA methylation for each cell type of placenta and corresponding maternal leukocyte samples or at single cell level that allows

researchers to investigate DNA methylation differences between different cell types and subtypes. Previous studies have demonstrated that the gene expression and DNA methylation profiles are different between cell types in placenta through single cell RNA sequencing [40, 41] and immunohistochemistry [42]. At present, single cell or cell type specific methylome for placenta were not well profiled, which highlights the need for further research focusing on DNA methylation profiles for individual cell types in placenta. While DNA methylation of cell types in circulating blood is well characterised [43, 44], and the cell proportion can be predicted through the DNA methylation profile of blood samples [45]. However, the DNA methylation difference between six types of circulating leukocytes from pregnant and non-pregnant women is not well characterised, which may aid biomarker identification for pregnancy health and deserves more research in future.

6.4 Conclusion

The work presented in this thesis in the field of DNA methylation analyses in placenta and maternal leukocytes aimed to profile genome-wide DNA methylation and biomarker identification. Through establishing a pipeline for EPIC array data analysis and exploration of the DNA methylation profile of human placenta and maternal leukocytes across early gestation, this work has contributed to the field in the following aspects. Firstly, this work demonstrated the importance of selecting appropriate methods for establishing analytic pipelines for data of interest which will benefit other future bioinformatics analyses based on EPIC array data. It also laid the foundation for further placenta research by showing DNA methylation changes throughout early placental development and showed a complex association between gene expression

and DNA methylation. Furthermore, this work showed that DNA methylation changes of maternal leukocytes may associate with maternal smoking status, maternal age and gestational age This adds to previous studies and informs future studies in biomarker identification for both pregnancy health and disease.

References

1. Mallik S, Odom GJ, Gao Z, Gomez L, Chen X, Wang L. An evaluation of supervised methods for identifying differentially methylated regions in Illumina methylation arrays. Briefings in bioinformatics. 2019;20(6):2224-35.
2. Maksimovic J, Phipson B, Oshlack A. A cross-package Bioconductor workflow for analysing methylation array data. F1000Research. 2016;5.
3. Müller F, Scherer M, Assenov Y, Lutsik P, Walter J, Lengauer T, et al. RnBeads 2.0: comprehensive analysis of DNA methylation data. Genome biology. 2019;20(1):55.
4. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics. 2017;33(24):3982-4.
5. Choufani S, Turinsky AL, Melamed N, Greenblatt E, Brudno M, Bérard A, et al. Impact of assisted reproduction, infertility, sex and paternal factors on the placental DNA methylome. Human molecular genetics. 2019;28(3):372-85.
6. Schroeder DI, Blair JD, Lott P, Yu HOK, Hong D, Crary F, et al. The human placenta methylome. Proceedings of the national academy of sciences. 2013;110(15):6037-42.
7. Monk D. Genomic imprinting in the human placenta. American journal of obstetrics and gynecology. 2015;213(4):S152-S62.
8. Li C, Fan Y, Li G, Xu X, Duan J, Li R, et al. DNA methylation reprogramming of functional elements during mammalian embryonic development. Cell discovery. 2018;4(1):1-12.

9. Greenberg MV, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. Nature reviews Molecular cell biology. 2019:1-18.

10. Walker N, Filis P, O'Shaughnessy PJ, Bellingham M, Fowler PA. Nutrient transporter expression in both the placenta and fetal liver are affected by maternal smoking. Placenta. 2019;78:10-7.

11. Förster R, Davalos-Misslitz AC, Rot A. CCR7 and its ligands: balancing immunity and tolerance. Nature Reviews Immunology. 2008;8(5):362-71.

12. Schanz A, Red-Horse K, Hess A, Baston-Büst D, Heiss C, Krüssel J. Oxygen regulates human cytotrophoblast migration by controlling chemokine and receptor expression. Placenta. 2014;35(12):1089-94.

13. Fernandez-Jimenez N, Allard C, Bouchard L, Perron P, Bustamante M, Bilbao JR, et al. Comparison of Illumina 450K and EPIC arrays in placental DNA methylation. Epigenetics. 2019;14(12):1177-82.

14. Pringle K, Kind K, Sferruzzi-Perri A, Thompson J, Roberts C. Beyond oxygen: complex regulation and activity of hypoxia inducible factors in pregnancy. Human reproduction update. 2010;16(4):415-31.

15. Brew O, Sullivan MH. Oxygen and tissue culture affect placental gene expression. Placenta. 2017;55:13-20.

16. Burton GJ, Jauniaux E, Murray AJ. Oxygen and placental development; parallels and differences with tumour biology. Placenta. 2017;56:14-8.

17. Salomon L, Sotiriadis A, Wulff C, Odibo A, Akolekar R. Risk of miscarriage following amniocentesis or chorionic villus sampling: systematic review of literature and updated meta-analysis. Ultrasound in Obstetrics & Gynecology. 2019;54(4):442-51.

18. Rahat B, Thakur S, Bagga R, Kaur J. Epigenetic regulation of STAT5A and its role as fetal DNA epigenetic marker during placental development and dysfunction. Placenta. 2016;44:46-53.

19. Harris AR, Nagy-Szakal D, Pedersen N, Opekun A, Bronsky J, Munkholm P, et al. Genome-wide peripheral blood leukocyte DNA methylation microarrays identified a single association with inflammatory bowel diseases. Inflammatory bowel diseases. 2012;18(12):2334-41.

20. Konwar C, Price EM, Wang LQ, Wilson SL, Terry J, Robinson WP. DNA methylation profiling of acute chorioamnionitis-associated placentas and fetal membranes: insights into epigenetic variation in spontaneous preterm births. Epigenetics & chromatin. 2018;11(1):1-14.

21. Delaney C, Garg SK, Yung R. Analysis of DNA methylation by pyrosequencing. Immunosenescence: Springer; 2015. p. 249-64.

22. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome biology. 2016;17(1):208.

23. Wang C, Chen M-H, Schifano E, Wu J, Yan J. Statistical methods and computing for big data. Statistics and its interface. 2016;9(4):399.

24. Morandat F, Hill B, Osvald L, Vitek J. Evaluating the design of the R language. In: European Conference on Object-Oriented Programming. 2012: Springer; p. 104-31.

25. Malviya A, Udhani A, Soni S. R-tool: Data analytic framework for big data. In: 2016 Symposium on Colossal Data Analysis and Networking (CDAN). 2016: IEEE; p. 1-5.

26. Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, et al. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome biology. 2014;15(2):R24.

27. White WM, Brost BC, Sun Z, Rose C, Craici I, Wagner SJ, et al. Normal early pregnancy: a transient state of epigenetic change favoring hypomethylation. Epigenetics. 2012;7(7):729-34.

28. Tekola-Ayele F, Workalemahu T, Gorfu G, Shrestha D, Tycko B, Wapner R, et al. Sex differences in the associations of placental epigenetic aging with fetal growth. Aging (Albany NY). 2019;11(15):5412.

29. Gonzalez TL, Sun T, Koeppel AF, Lee B, Wang ET, Farber CR, et al. Sex differences in the late first trimester human placenta transcriptome. Biology of sex differences. 2018;9(1):4.

30. Avila L, Yuen R, Diego-Alvarez D, Penaherrera M, Jiang R, Robinson W. Evaluating DNA methylation and gene expression variability in the human term placenta. Placenta. 2010;31(12):1070-7.

31. Teschendorff AE, Zheng SC. Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. Epigenomics. 2017;9(5):757-68.

32. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. The Lancet. 2018;392(10149):777-86.

33. Ma B, Allard C, Bouchard L, Perron P, Mittleman MA, Hivert M-F, et al. Locus-specific DNA methylation prediction in cord blood and placenta. Epigenetics. 2019;14(4):405-20.

34. Assunção MD, Calheiros RN, Bianchi S, Netto MA, Buyya R. Big Data computing and clouds: Trends and future directions. Journal of Parallel and Distributed Computing. 2015;79:3-15.

35. Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS computational biology. 2017;13(11):e1005752.

36. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics. 2008;9(1):559.

37. Tremblay BL, Guénard F, Lamarche B, Pérusse L, Vohl M-C. Network Analysis of the Potential Role of DNA Methylation in the Relationship between Plasma Carotenoids and Lipid Profile. Nutrients. 2019;11(6):1265.

38. Horvath S. DNA methylation age of human tissues and cell types. Genome biology. 2013;14(10):3156.

39. Lee Y, Choufani S, Weksberg R, Wilson SL, Yuan V, Burt A, et al. Placental epigenetic clocks: estimating gestational age using placental DNA methylation levels. Aging (Albany NY). 2019;11(12):4238.

40. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, et al. Single-cell reconstruction of the early maternal–fetal interface in humans. Nature. 2018;563(7731):347-53.

41. Liu Y, Fan X, Wang R, Lu X, Dang Y-L, Wang H, et al. Single-cell RNA-seq reveals the diversity of trophoblast subtypes and patterns of differentiation in the human placenta. Cell research. 2018;28(8):819-32.

42. Wilson RL, François M, Jankovic-Karasoulos T, McAninch D, McCullough D, Leifert WR, et al. Characterization of 5-methylcytosine and 5-

hydroxymethylcytosine in human placenta cell types across gestation. Epigenetics. 2019;14(7):660-71.

43. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlén S-E, Greco D, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PloS one. 2012;7(7):e41361.

44. Zilbauer M, Rayner TF, Clark C, Coffey AJ, Joyce CJ, Palta P, et al. Genome-wide methylation analyses of primary human leukocyte subsets identifies functionally important cell-type–specific hypomethylated regions. Blood. 2013;122(25):e52-e60.

45. Salas LA, Koestler DC, Butler RA, Hansen HM, Wiencke JK, Kelsey KT, et al. An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. Genome biology. 2018;19(1):1-14.