VNIVERSITAT DE VALÈNCIA

PhD program in Biomedicine and Biotechnology

Epigenetic analysis of cell-free DNA as a tool to study pathogenesis and maternal syndrome of preeclampsia

Author:

Irene Corachán García

Supervised by:

Dr. Antonio Diez Juan

Dr. Carlos Simón Vallés

Dr. Felipe Vilella Mitjana

Valencia, October 2019

El Dr. Antonio Diez Juan, Doctor en Biología, y mánager del departamento de Desarrollo de Productos en Igenomix; El Dr. Carlos Simón Vallés, catedrático de Pediatría, Obstetricia y Ginecología de la Universidad de Valencia y director Científico de Igenomix y el Dr. Felipe Vilella Mitjana, director del departamento de investigación de Igenomix Foundation e investigador Miguel Servet del Incliva.

CERTIFICAN:

Que el trabajo de investigación titulado: "Epigenetic analysis of cell-free DNA as a tool to study pathogenesis and maternal syndrome of preeclampsia" ha sido realizado íntegramente por Irene Corachán García bajo su dirección. Dicha memoria está concluida y reúne todos los requisitos para su presentación y defensa como TESIS DOCTORAL ante un tribunal.

Y para que así conste a los efectos oportunos, firman la presente certificación en Valencia a 11 de Octubre de 2019.

Fdo. Dr. Antonio Diez Juan

Fdo. Dr. Carlos Simón Vallés

Fdo. Dr. Felipe Villela Mitjana

AGRADECIMIENTOS

Parecía que este momento no iba a llegar nunca, pero aquí estoy, por fin, escribiendo los agradecimientos. En primer lugar, me gustaría agradecer a mis tres directores su labor en esta tesis. En especial, al Dr. Carlos Simón, gracias por darme la oportunidad de trabajar en un lugar como Igenomix. Al Dr. Antonio Diez, por todo lo que has hecho para que esta tesis funcione. Sabes que sin ti no habría podido hacerlo, gracias por aguantarme en mis momentos más *dreamkiller.*

También me gustaría agradecer esta tesis a mis compañeros. En este momento es inevitable acordarse del comienzo de todo, y ahí aparecen Sebas y Pablo. Recuerdo el primer día, cuando me sentaron entre vosotros… gracias a los dos por acogerme desde el minuto uno, por la complicidad, por haberme tratado siempre tan bien y por todos los momentos compartidos. Gracias Pablo, por hacer de Mas de Borràs un lugar inolvidable. Gracias Sebas, por creer siempre en mi. Para completar el equipo estaba David; gracias, por aportar tu cordura en los momentos de más locura.

Después llegó la mudanza, y con ella un nuevo compañero a mi lado. Gracias Gus, por haberme apoyado y defendido siempre, pero, sobre todo, por enseñarme *The Noun Project*.
Más tarde se unió Begoña, qué pena que no llegaras antes… Gracias por escucharme siempre que lo he necesitado.
A todos vosotros, *A-pollo team*, quiero agradeceros que la comida del viernes se convirtiera en el mejor momento de la semana.

En estos agradecimientos tienen una parte muy importante dos personas que me han alegrado los días y ayudado a partes iguales. Gracias Ale, por tu dulzura, el mundo sería mucho mejor si hubiera más gente como tú. Gracias Iolanda, por aguantar todos mis agobios y por estar siempre dispuesta a ayudarme. Has sido mi gran apoyo en estos últimos meses. Os quiero mucho.

También me gustaría destacar a dos personas que me han ayudado siempre que lo he necesitado con una paciencia infinita. Gracias Jorge, por "empujarme" a ese mundo que tanto me asustaba y que terminó gustándome. Gracias Roser, por explicarme todo tan bien, hasta en la distancia, y por venir a visitarme a mi sitio y hacerme reír.

Por otro lado, me gustaría agradecer a todos mis compañeros de investigación, Pati, Marta, Tere, Fany, David, Kaisu, Alba, por su ayuda cuando lo he necesitado, pero también por ser mis compañeros en los mejores momentos del día: almuerzos, comidas, *oliva days*… Gracias Nuria, por proximidad te tocó aguantarme y siempre que te he pedido ayuda me la has dado. Gracias Ana, por alegrarme los días con nuestras conversaciones.

Gracias a toda la gente del laboratorio que me ha ayudado, sobre todo a aquellos que lo han hecho con una sonrisa. En especial a las chicas NACE y CGT porque, desde los comienzos, siempre que os he necesitado habéis estado ahí.

Por otro lado, me gustaría agradecer a la gente que me acompaña en mi vida y que ha tenido que sufrir, en mayor o menor medida, esta tesis.

A mis amigas, Bárbara, Cris, Ester, Nuri y Patri, gracias por todas las risas, con vosotras cualquier plan es el mejor plan.

A mi familia. Cuando pienso en vosotros es inevitable sentir gratitud, y darme cuenta de lo afortunada que soy por teneros. Para mi resulta fácil, pero no siempre se tiene una familia en la que confiar plenamente y sentirse querido. Quiero daros las gracias por tanto… A mi abuela, por alegrarte más que yo de las cosas buenas que me pasan. A mis tías, Amparín, Lupe Y Mª José, por seguir mimándome tanto por mucho que pasen los años. A mis primas, Belén y Amparo, por todas nuestras tonterías y risas, que espero que nunca se acaben por mucho que crezcamos. A mis primos, Vega, Mateo y Candela, por hacer que se me olvide todo cuando estoy con vosotros. A mis hermanas, no podría imaginarme una vida sin vosotras, sabéis cuánto os quiero y os necesito. Gracias por aguantarme siempre. A mi padre, por enseñarme a coger una hoja en blanco cuando algo no me salía y volverlo a intentar. A mi madre, por enseñarme que lo más importante en esta vida es ser buena persona y, sobre todo, por tu amor incondicional. A Migue, por cuidarnos tanto a todas. A mis suegros, cuñados y sobrinos, por hacerme sentir siempre una más en vuestra familia. A Andrés, por quererme hasta cuando no me quiero ni yo.

El presente trabajo de tesis doctoral ha sido realizado en los laboratorios de la Fundación Igenomix. Gracias a la ayuda para la contratación de personal investigador en formación de carácter predoctoral, Programa VALi+d de la Generalitat Valenciana y el Fondo Social Europeo (ACIF/2016/024).

RESUMEN

INTRODUCCIÓN:

Trastornos hipertensivos del embarazo: Preeclampsia

Alrededor del 10% de los embarazos en todo el mundo sufren trastornos hipertensivos. La *Task Force on Hypertension in Pregnancy* del Colegio Americano de Ginecólogos y Obstetras (ACOG) estableció una clasificación de estos, diviéndolos en cuatro categorías: hipertensión crónica, hipertensión gestacional, preeclampsia-eclampsia y preeclampsia sobreañadida a hipertensión crónica (ACOG, 2013).

La hipertensión crónica se define por la presencia de presión sanguínea elevada antes de la gestación o que aparece antes de las 20 semanas de la misma. La hipertensión gestacional ocurre cuando la hipertensión aparece después de las 20 semanas, sin ir acompañada de proteinuria, y desaparece después del parto. La preeclampsia se define como la aparición de hipertensión después de las 20 semanas de gestación, acompañada por la aparición de proteinuria y/o síntomas multi-sistémicos en la madre (ACOG, 2013). La preeclampsia puede verse complicada con eclampsia debido a la aparición de convulsiones de tipo gran mal (Phipps *et al.*, 2019).

De todos ellos, la preeclampsia es el más común, afectando a entre un 2 y 8% de los embarazos en todo el mundo (Duley, 2009). Además es, junto con la eclampsia, una de las 3 principales causas de mortalidad y morbilidad materna (Ghulmiyyah and Sibai, 2012). Actualmente, el diagnóstico de la preeclampsia según la clasificación establecida por ACOG se basa en la elevación de la presión sanguínea (sistólica ≥ 140mmHg y/o diastólica ≥ 90 mmHg) después de la semana 20 de gestación, junto con la aparición de proteinuria o, en ausencia de esta, hipertensión acompañada de: trombocitopenia, insuficiencia renal, edema pulmonar, alteración de la función hepática o alteraciones cerebrales o visuales. La preeclampsia grave ocurre cuando los niveles de presión sanguínea alcanzan sistólica ≥ 160mmHg y/o diastólica ≥ 110 mmHg o las complicaciones sistémicas empeoran considerablemente (ACOG, 2013).

Patogénesis de la preeclampsia

La patogénesis de la preeclampsia se divide en 2 etapas: placentación defectuosa y desarrollo del síndrome materno. Brevemente, una placentación anormal provoca una isquemia en la placenta, la cual libera a la circulación materna una serie de factores que provocan una disfunción endotelial y los signos maternos del síndrome (Phipps *et al.*, 2019; Chaiworapongsa *et al.*, 2014; Redman and Sargent, 2005).

Durante el proceso de placentación, las arterias espirales maternas sufren un aumento de calibre que garantiza una correcta perfusión placentaria. En la preeclampsia, la transformación de las arterias espirales sólo ocurre en los segmentos superficiales, presentes en la decidua, mientras que los segmentos presentes en el miometrio no aumentan su calibre (Young *et al.*, 2010). Como resultado, el flujo sanguíneo es insuficiente provocando una isquemia en la placenta, así como una posterior restricción del crecimiento intrauterino. La causa por la que ocurre este fallo en la transformación de las arterias espirales maternas es desconocida hoy en día. Entre las causas propuestas se encuentran fallos en la regulación de la diferenciación de los trofoblastos mediada por la hipoxia (Caniggia *et al.*, 2000; Caniggia *et al.*, 1999; Rajakumar *et al.*, 2004) o fallos en el sistema inmune durante el reconocimiento materno-fetal (Young *et al.*, 2010; Chaiworapongsa *et al.*, 2014).

La segunda etapa de la patogénesis de la preeclampsia se define como la aparición del síndrome materno. Este síndrome se caracteriza por una respuesta inflamatoria sistémica, disfunción endotelial y la aparición de hipertensión, proteinuria, alteraciones visuales, edema pulmonar y fallos renales y hepáticos. Se ha propuesto que el síndrome es causado por los factores liberados al torrente sanguíneo por la placenta. Entre otros, se conoce que el desequilibrio en los factores angiogénicos/antiangiogénicos presentes en la circulación materna tiene un papel importante en la preeclampsia (Redman and Sargent, 2005; Young *et al.*, 2010).

A pesar de la gran relevancia de este síndrome, su fisiopatología es todavía desconocida, por lo que hoy en día muchos trabajos se centran en mejorar el conocimiento de la misma. La existencia de diferencias de metilación asociadas a enfermedades (Robertson, 2005; Jin and Liu, 2018; (Fernandez *et al.*, 2012) sugiere que el estudio de la metilación del ADN podría

suponer un enfoque diferente en el estudio de la preeclampsia, pudiendo proporcionar información relevante para su fisiopatología.

Epigenética: Metilación del ADN

La epigenética se define como el estudio de toda aquella información presente en el ADN, distinta de la propia presente en la secuencia de nucleótidos, que tiene un efecto en la expresión génica. Entre los cambios epigenéticos se incluyen modificaciones de la cromatina, así como modificaciones del ADN (Greally, 2018). La metilación del ADN es, sin lugar a duda, la modificación epigénetica más estudiada hasta la fecha (Smith and Meissner, 2013). Consiste en la adición de un grupo metilo en una citosina, y es un proceso dinámico ya que puede ocurrir de novo, mantenerse y borrarse (Chen *et al.*, 2017). La metilación sólo puede ocurrir en citosinas, y normalmente ocurre en citosinas seguidas de una guanina, conocidos como sitios CpG. El genoma humano contiene alrededor de 28 millones de estos sitios, lo que representa menos de un 1% del genoma (Smith and Meissner, 2013). De ellos, una parte se encuentra concentrada en unas regiones conocidas como islas CpG. Se calcula que existen unas 3000 en el genoma humano y se caracterizan por el enriquecimiento en sitios CpG. Mientras que los sitios CpG fuera de las islas se encuentran metilados en un 60-80% de los casos, las islas CpG suelen permanecer no metiladas, sobre todo cuando se encuentran asociadas a regiones promotoras (Bergman and Cedar, 2013).

La metilación del ADN tiene un papel fundamental en procesos tan importantes como el desarrollo embrionario o la gametogénesis (Dor and Cedar, 2018; Cedar and Bergman, 2012; von Meyenn and Reik, 2015; Barlow and Bartolomei, 2014). Además, su función como mecanismo regulatorio de la expresión génica ha sido ampliamente estudiada. Aunque tradicionalmente la metilación se ha considerado un mecanismo de silenciamiento de genes, hoy en día se piensa que el efecto en la expresión depende de la posición del genoma donde ocurra (Jones, 2012).

Enfermedad y metilación

Debido a la gran importancia de la metilación del ADN en el desarrollo, así como en la regulación génica, su papel durante la enfermedad ha sido estudiado y cada vez más trabajos evidencian la implicación de la metilación en el desarrollo y progresión de distintas enfermedades en humanos (Robertson, 2005; Jin and Liu, 2018). Aunque se han descrito

cambios en la metilación del ADN asociados a enfermedades no cancerosas como el alzheimer o enfermedades autoinmunes (Fernandez *et al.*, 2012), la mayoría de los trabajos se han centrado en estudiar su implicación en el cáncer. Gracias a ello se ha definido un patrón de metilación común en las células tumorales caracterizado por la pérdida global de metilación en sitios CpG fuera de las islas CpG, junto con una hipermetilación de estas. Se ha propuesto que la pérdida de metilación conllevaría un aumento de la inestabilidad genómica, mientras que el aumento en la hipermetilación de islas CpG provocaría un silenciamiento de genes supresores de tumores (Robertson, 2005; Baylin and Jones, 2016).

Preeclampsia y metilación

Del mismo modo que el papel de la metilación del ADN se ha estudiado en otras enfermedades en busca de nueva información relevante para el conocimiento de la misma, los cambios en la metilación también se han estudiado en la preeclampsia. Como resultado, la mayoría de los trabajos describen la presencia de diferencias de metilación en el ADN de la placenta entre mujeres con preeclampsia y embarazos controles (resumidos en Tablas 2 y 3). Estos cambios también han sido descritos en células sanguíneas tanto maternas como fetales, células del cordón umbilical y arterias maternas. A pesar de la presencia de cambios en la metilación asociados a preeclampsia, la gran heterogeneidad entre los distintos estudios no permite establecer conclusiones claras de todos ellos.

ADN libre circulante

De un tiempo a esta parte, el estudio del ADN libre circulante (ADNlc) ha surgido como una valiosa fuente de información en el campo de la biomedicina como una posibilidad para estudiar el estado global de un individuo de una manera mínimamente invasiva. El ADNlc es el conjunto de moléculas de ADN que se encuentran en la circulación sanguínea libres de células. Estas moléculas son una mezcla de ADN proveniente de distintas partes del cuerpo, siendo la mayor parte de ellas de las células hematopoyéticas (Moss *et al.*, 2018; Sun *et al.*, 2015). Durante el embarazo, a esta mezcla de moléculas se unen las provenientes de la placenta (Bianchi *et al.*, 2015), así como durante los procesos cancerosos se unen moléculas provenientes del tumor (Vasioukhin *et al.*, 1994; Sorenson *et al.*, 1994). La presencia de estas moléculas se utiliza hoy en día para el cribado prenatal no invasivo y la detección y monitorización del cáncer, respectivamente (Corcoran and Chabner, 2018; Bianchi *et al.*,

2015). El tamaño de las moléculas presentes en el ADNlc depende de su origen. En el caso del ADN de origen mitocondrial, presente en menor medida que el nuclear, la mayoría de moléculas tienen un tamaño ~ 100pb (Zhang *et al.*, 2016). Por otro lado, la mayoría de moléculas de ADN de origen nuclear tienen un tamaño ~ 166 pb, que corresponde con el tamaño del ADN enrollado alrededor de un nucleosoma (~147 pb) sumado al ADN espaciador asociado a la histona H1 (Wan *et al.*, 2017). A este pico donde se encuentran la mayoría de moléculas le siguen dos picos de menor altura correspondientes al ADN empaquetado en forma de 2 y 3 nucleosomas, respectivamente (Volik *et al.*, 2016).. Debido a que la mayor parte de moléculas tienen un origen nuclear, este perfil de tamaños es el característico del ADN libre circulante. Existen también diferencias en función del tejido de origen debidas principalmente a las diferencias en la compactación de la cromatina entre tejidos (Snyder *et al.*, 2016; Sun *et al.*, 2018a). ; el ADN derivado del núcleo de células no hematopoyéticas es generalmente más pequeño que el derivado de las células hematopoyéticas (Zheng *et al.*, 2012).

La vida media de estas moléculas es de aproximadamente 1-2 horas, y en el caso de el embarazo (Kustanovich *et al.*, 2019), el ADN de origen placentario desaparece a las 24-48 horas del parto (Yu *et al.*, 2013). Se ha propuesto que el origen de las moléculas es mayoritariamente la apoptosis celular, y el hecho de que el tamaño de estas corresponda con el obtenido por la digestión por proteínas de la familia de las caspasas apoya esta hipótesis (Kustanovich *et al.*, 2019).

ADN libre circulante y epigenética

Además de la información contenida en la secuencia de ADN, el ADNlc contiene información epigenética como modificaciones de histonas o metilación. El hecho de que estas modificaciones puedas ser estudiadas en el ADNlc supone una valiosa fuente de información en la búsqueda de marcadores en condiciones que no impliquen cambios en la secuencia de ADN. La existencia de patrones de metilación específicos de tejido (Fernandez *et al.*, 2012; Kundaje *et al.*, 2015) ha sido utilizada para estudiar el tejido de origen de las moléculas presentes en el ADNlc (Sun *et al.*, 2015; Lehmann-Werman *et al.*, 2016; Moss *et al.*, 2018). En enfermedades como el cáncer, en las cuales se produce un aumento en la cantidad de ADNlc proveniente del tumor, conocer el origen de las moléculas resulta de especial utilidad. Además, también ha sido empleado para la monitorización en trasplantes de órganos o en

otras enfermedades como la diabetes tipo 1 o en casos de sepsis (Lehmann-Werman *et al.*, 2016; Moss *et al.*, 2018). A parte de estas situaciones que implican un aumento del ADNlc proveniente de un tejido, muchos estudios se han centrado en la búsqueda de regiones diferencialmente metiladas en el ADNlc entre individuos sanos y enfermos. De hecho, actualmente ya existen marcadores basados en cambios en la metilación del ADNlc para determinados tipos de cáncer (Feng *et al.*, 2018; Gai and Sun, 2019).

Otra fuente de información contenida en el ADNlc es la proporcionada por los patrones de fragmentación de las moléculas. Estudios recientes se han enfocado en el estudio de los tamaños de los fragmentos de ADN presentes en la circulación. En el caso del cáncer y de las mujeres embarazadas la presencia de ADNlc proveniente del tumor y de la placenta, respectivamente, implica un aumento de moléculas con un tamaño menor a 166pb (Mouliere *et al.*, 2018; Jiang *et al.*, 2015; Hellwig *et al.*, 2018; Lo *et al.*, 2010). Esta diferencia está siendo utilizada para mejorar la detección de ADNlc tumoral, así como para enriquecer en ADNlc fetal en el cribado prenatal no invasivo (Hellwig *et al.*, 2018; Mouliere *et al.*, 2018).

ADN libre circulante y preeclampsia

En este campo, los estudios llevados a cabo en preeclampsia han estado enfocados únicamente en la medición de la concentración de ADNlc y en su utilización como marcador predictivo de la enfermedad. Los estudios realizados hasta la fecha relacionan la preeclampsia con un aumento en los niveles absolutos de ADNlc fetal (Lo *et al.*, 1999; Martin *et al.*, 2014; Contro *et al.*, 2017). Sin embargo, aquellos que analizan la proporción de ADNlc de origen fetal con respecto al total, conocido como fracción fetal, encuentran una disminución de la misma en los casos de preeclampsia (Rolnik *et al.*, 2015; Rolnik *et al.*, 2018; Bender *et al.*, 2019; Gerson *et al.*, 2019). Diversos estudios han propuesto la utilización de la baja fracción fetal como un marcador de riesgo para el desarrollo de preeclampsia (Rolnik *et al.*, 2018; Gerson *et al.*, 2019). Sin embargo, su utilización como marcador independiente requiere más investigación.

HIPÓTESIS:

La hipótesis de este trabajo es que existen regiones diferencialmente metiladas en el ADN libre circulante entre pacientes con preeclampsia grave y embarazos control. La

detección de tales diferencias facilitaría el estudio para definir su fisiopatología y el descubrimiento de nuevos biomarcadores para su diagnóstico y pronóstico.

OBJETIVOS:

- Establecer un método para la detección de cambios en la metilación del ADN libre circulante aplicable a una rutina de laboratorio de diagnóstico clínico.
- Comparar el metiloma del ADN libre circulante entre mujeres embarazadas con preeclampsia grave y sin ella en busca de regiones diferencialmente metiladas.
- Explorar el papel de los cambios de metilación en la fisiopatología y síntomas de la preeclampsia grave.
- Investigar la utilidad de otras características del ADN libre circulante, como los patrones de fragmentación, en la identificación de biomarcadores asociados a la preeclampsia.

METODOLOGÍA:

Con el fin de estudiar las diferencias de metilación en ADNlc asociadas a preeclampsia, se tomaron 41 muestras de sangre a mujeres diagnosticadas con preeclampsia grave (del inglés, sPE), así como mujeres embarazadas sin preeclampsia. El estudio fue aprobado por el Comité de Ética del Hospital Univeristario y Politécnico La Fe (Valencia, España) (2011/0383) y todas las mujeres incluidas en el estudio firmaron el consentimiento informado.

Procesamiento de las muestras y obtención de ADN libre circulante

La muestra de sangre periférica fue recogida en tubos cell-free DNA BCT (Streck, NE, USA). Para la obtención del plasma, se llevó a cabo una primera centrifugación a 4ºC, 1600 rcf durante 15 minutos, el sobrenadante fue sometido a una segunda centrifugación a 4ºC, 16000 rcf durante 10 minutos. El sobrenadante resultante (1mL) se usó para obtener el ADNlc utilizando para ello un protocolo modificado de QIAsymphony DSP Virus/Pathogen Kit (QIAGEN, Hilden, Germany) se llevó a cabo usando el sistema QIAsymphony SP. Con el objetivo de mejorar el rendimiento final se utilizó carrier RNA y el ADNlc fue eluido en un volumen final de 50 μL.

Obtención de fracción metilada

A partir de los 50 µL de ADNlc, la fracción de ADN metilada fue obtenida utilizando el kit MethylMinerTM Methylated DNA Enrichment Kit (Invitrogen, Carlsbad, USA). Este kit utiliza bolas magnéticas con streptavidina, a las cuales se une el dominio de unión metilcitosina de la proteína MBD2 que se encuentra biotinilado (enriquecimiento MBD). El ADN es incubado con las bolas magnéticas y aquellas moléculas que contengan citosinas metiladas se unirán a las bolas y serán capturadas. El protocolo se llevó a cabo siguiendo las instrucciones del fabricante. Los volúmenes utilizados por muestra fueron: 2.5 µL de Dynabeads M-280 Streptavidin y 1.5 µL de proteína. El proceso de elución se llevó a cabo en 3 pasos: el primero, con un tampón de baja concentración de sal (0 mM NaCl) para eliminar moléculas de ADN no metiladas; el segundo con un tampón 450 mM NaCl, para eliminar la posible presencia de restos de ADN no metilados, así como moléculas de ADN con baja presencia de sitios CpG metilados. Por último, la fracción metilada fue obtenida tras el tercer y último paso de elución, que se llevó a cabo utilizando un tampón 2000mM NaCl. Con el fin de eliminar restos del tampón utilizado, los 200 µL de ADN metilado fueron purificados usando 200 µL de bolas magnéticas Agencourt AMPure XP (Beckman Coulter, Brea, USA) siguiendo las instrucciones del fabricante. Finalmente, el ADN fue eluido en un volumen de 50 µL.

Preparación de librería y secuenciación

A partir de los 50 µL de ADNlc metilado se llevó a cabo la preparación de librerías utilizando el kit TruSeq Nano DNA Library Prep Kit (Illumina Inc., San Diego, USA) siguiendo las instrucciones del fabricante. Brevemente, los extremos de las moléculas de ADN fueron convertidos en extremos romos, ligando una adenina a los extremos 3’ de las moléculas para evitar que los fragmentos se unan entre si en el siguiente paso; la ligación de los adaptadores de secuenciación. Una vez que los adaptadores se ligaron a las moléculas de ADNlc, se llevó a cabo una purificación para eliminar los posibles restos de adaptadores. Seguidamente los fragmentos fueron amplificados mediante PCR, incluyendo la adición de un índice único para cada muestra. El ADN resultante fue purificado y cuantificado.

Se llevaron a cabo dos carreras de secuenciación, ambas utilizando el secuenciador NextSeq 500 sequencing system (Illumina Inc., San Diego, USA). La primera de las carreras fue del tipo *single-read,* en la cual las moléculas son secuenciadas sólo por un extremo (150 ciclos) y se

llevó a cabo utilizando 24 muestras (12 por grupo). La segunda carrera fue del tipo *paired-end*, en la cual las moléculas son secuenciadas por ambos extremos (2 x 100 ciclos), y se llevó a cabo utilizando 41 muestras (Control n=22, sPE n=19).

Análisis bioinformático de los datos:

A no ser que se especifique lo contrario, los análisis resumidos a continuación se llevaron a cabo de igual forma para los datos obtenidos a partir de las dos carreras de secuenciación.

- **Calidad y Alineamiento**

Los datos fueron descargados del secuenciador utilizando Illumina Basespace, se obtuvieron los archivos FASTQ utilizando bcl2fastq (Illumina) para llevar a cabo un análisis de la calidad de secuenciación. Seguidamente, las secuencias fueron alineadas con el genoma de referencia (hg19) utilizando BWA. Las lecturas con una calidad de mapeo menor de 10, así como los duplicados encontrados fueron eliminados utilizando PICARD (https://broadinstitute.github.io/picard/index.html).

- **Cuantificación del cromosoma Y**

Con el fin de llevar a cabo una estimación de la contribución fetal al ADNlc los datos de secuenciación de todas las muestras fueron utilizados para seleccionar aquellas regiones del cromosoma Y con baja homología con regiones del cromosoma X. Seguidamente, se estimó la contribución fetal al ADNlc en aquellos embarazos con feto masculino. Para ello, se calculó la cobertura media normalizada (Cuentas por Millón, CPM) de las regiones obtenidas previamente regiones para aquellas muestras correspondientes a embarazados de un feto masculino.

- **Regiones Diferencialmente Metiladas**

Para la detección de regiones diferencialmente metiladas (del inglés, DMRs), se utilizó el software Methylaction (http://github.com/jeffbhasin/methylaction). Una vez obtenidas las DMRs, fueron anotadas utilizando el software Goldmine (https://github.com/jeffbhasin/goldmine). La anotación se llevó a cabo para contexto genómico, islas CpG Y sitios de unión de factores de transcripción (del inglés, TFBS).

Para el análisis posterior, las regiones fueron filtradas (pvalue < 0.01 y |log2(fold change)| > log2(1)) y divididas en regiones hipermetiladas e hipometiladas en sPE en comparación con control.

- Análisis funcional:

Las regiones anotadas como islas CpG y promotores fueron obtenidas para ambos grupos de DMRs. El análisis funcional para el enriquecimiento en rutas KEGG se llevó a cabo utilizando la lista de genes correspondientes a esas regiones utilizando el Molecular Signatures Database v6.2 (MSigDB) con el software GSEA (http://software.broadinstitute.org/gsea/msigdb/). Las 10 rutas más significativamente enriquecidas fueron obtenidas (FDR < 0.05). Por otro lado, se seleccionaron aquellos TFBS que aparecían sobre-representados en cada grupo de DMRs. Se obtuvieron los factores de transcripción correspondientes a los sitios sobre-representados en regiones hiper- e hipo-metiladas en sPE. Los factores de transcripción se analizaron utilizando Overrepresentation enrichment Analysis (ORA) de WebGestalt (WEB-based GEne SeT AnaLysis Toolkit) (Liao *et al.*, 2019)), como resultado se obtuvieron los 10 procesos biológicos más significativamente enriquecidos.

- DMRs y enfermedad:

Las regiones anotadas como promotores, que tenían una buena cobertura en ambos grupos de muestras fueron utilizadas para analizar la presencia de genes relacionados con la patología. Para ello, en primer lugar, se llevó a cabo un análisis de asociación gen-enfermedad. La herramienta online DAVID se utilizó para estudiar la presencia de genes relacionados con alguna enfermedad usando para ello la base de datos GAD (del inglés, Genetic Association Database), se obtuvieron las 5 enfermedades más significativamente asociadas con los genes. Adicionalmente, con el fin de realizar un análisis más enfocado a la preeclampsia, se obtuvieron genes previamente relacionados con la misma utilizando referencias bibliográficas, así como los datos presentes en la base de datos de Ingenuity Pathway Analysis (QIAGEN). La lista de genes presentes en la DMRs se comparó con la lista de genes relacionados con preeclampsia.

- **Análisis de los tamaños de fragmentos**

En análisis basado en los tamaños de fragmentos del ADNlc sólo se llevó a cabo con los datos procedentes de la carrera de secuenciación *paired-end* ya que la información del tamaño de fragmento no puede extraerse a partir de otro tipo de datos.

Se utilizó la función CollectInsertSizeMetrics de PICARD para obtener los datos de tamaño de inserto a partir de los archivos bam de cada muestra. La proporción de lecturas para distintos rangos de tamaños se calculó dividiendo el número de lecturas con un tamaño correspondiente al establecido en el rango por el total de lecturas de esa muestra. Utilizando los datos de estimación fetal obtenidos previamente, se analizó la correlación entre la proporción de fragmentos y la fracción fetal.

- **Análisis estadístico**

Todos los análisis se realizaron utilizando los programas estadísticos R (www.r-project.org) y GraphPad Prism (La Jolla, CA). El programa R fue utilizado para testar la normalidad de los datos (saphiro.test), para comparar las proporciones de DMRs anotadas en los distintos contextos (prop.test) y para representar la distribución de fragmentos. También utilizando R se estudió la correlación entre proporción de fragmentos y fracción fetal usando para ello el paquete ggpubr. GraphPad Prism se utilizó para representar las proporciones de anotación, para representar y analizar la cuantificación del cromosoma Y, así como la proporción de lecturas en distintos rangos de tamaño de fragmentos. La prueba D'Agostino-Pearson omnibus se utilizó para testar la normalidad de los datos y en base al resultado se utilizó t-test no pareado o la prueba de Mann-Whitney. Los p-valores ≤ 0.05 se consideraron estadísticamente significativos.

RESULTADOS:

- Regiones diferencialmente metiladas:

Los resultados obtenidos en las dos carreras de secuenciación fueron muy similares. Se observó un enriquecimiento en promotores entre las regiones hipermetiladas en sPE, así como un enriquecimiento en regiones intergénicas e intrónicas en las regiones hipometiladas. Además, la presencia de regiones en islas CpG fue significativamente superior en las regiones hipermetiladas en sPE, siendo la mayoría regiones promotoras. El análisis funcional mostró

cambios de metilación en genes implicados en rutas de señalización como la ruta de señalización Wnt, la ruta de señalización del VEGF, la ruta de contracción del músculo liso vascular o rutas de señalización en cáncer. Además, los factores de transcripción correspondientes a sitios sobre-representados en las regiones hipermetiladas estaban asociados con una regulación positiva de procesos biológicos, mientras que lo contrario se observó para aquellos correspondientes a sitios más representadas en las regiones hipometiladas. Se encontraron asociaciones entre genes con promotores hipermetilados y genes implicados en algunos de los síntomas (Hipertensión, enfermedades renales) y en las secuelas a largo plazo (Diabetes Tipo 2, niveles de colesterol HDL, angina coronaria). Entre los genes con promotores hipermetilados encontramos algunos cuya implicación en la preeclampsia ha sido previamente descrita como endoglina (*ENG*), factor de crecimiento endotelial vascular B (*VEGFB)* u oxido nitro sintasa endotelial (*eNOS)*.

- Patrones de fragmentación del ADNlc:

Los resultados obtenidos para los datos de la carrera de secuenciación *paired-end* muestran un enriquecimiento en fragmentos pequeños asociados a los embarazos con preeclampsia grave. Estas diferencias se observan tanto en los tamaños correspondientes al pico de 1 nucleosoma, como al de 2 nucleosomas. El análisis de correlación mostró que existe una correlación negativa entre los fragmentos “pequeños” (270-320 pb) y la fracción fetal, así como una correlación positiva entre los fragmentos “grandes” (175-225 pb) y la misma.

DISCUSIÓN:

El estudio del ADNlc ha surgido como una valiosa fuente de información ya que es una mezcla de fragmentos de ADN de distintas partes del cuerpo. En campos con el cáncer, la biopsia líquida es ya una realidad, habiéndose demostrado su uso en su pronóstico y monitorización (Wan *et al.*, 2017). Además de en el cáncer, el otro gran uso del ADNlc actualmente es en el cribado prenatal no invasivo (Gai and Sun, 2019). Sin embargo, su uso no está muy extendido en otros campos. El estudio de las marcas epigenéticas representa una oportunidad para ampliar su aplicación a otras situaciones. Trabajos recientes han conseguido identificar el tejido de origen de los fragmentos de ADNlc mediante el uso de marcas epigenéticas (Snyder *et al.*, 2016; Lehmann-Werman *et al.*, 2016; Moss *et al.*, 2018). Además, estas se han utilizado para la identificación de cambios entre individuos sanos y enfermos (Fernandez *et al.*, 2012).

Así, esta tesis parte de la premisa de que los cambios en la metilación del ADNlc podrían emplearse para la búsqueda de marcadores de patología, así como de las complicaciones asociadas a la misma.

Actualmente existen numerosos métodos para el estudio de la metilación del ADN, en este trabajo se ha llevado a cabo la secuenciación de ADN metilado enriquecido utilizando el dominio de unión a la metilcitosina (del inglés, MBD-seq). Los métodos de enriquecimiento del ADN metilado suponen una alternativa perfecta a la secuenciación bisulfito en los casos en los que existe baja concentración de muestra dado que no se produce pérdida de ADN durante el proceso. A pesar de que los métodos de enriquecimiento del ADN metilado no proporcionan una resolución a nivel de base única, su capacidad para la detección de regiones diferencialmente metiladas (del inglés, DMRs) ha sido demostrada. Por lo tanto, la secuenciación MBD-seq supone un enfoque rentable para el análisis de cambios asociados a la enfermedad. Además, en un contexto clínico, representa una opción asequible, que permitiría el análisis de un número elevado de muestras (Harris *et al.*, 2010).

En este trabajo, se llevó a cabo una primera carrera de secuenciación del tipo *single-read*, como un primer paso económico y rápido para el descubrimiento de cambios asociados a enfermedad. Seguidamente, una carrera de secuenciación *paired-end* se utilizó para obtener un mayor conocimiento de las regiones, así como para estudiar los patrones de fragmentación del ADNlc. Estudios recientes han demostrado la utilidad del análisis de los tamaños de las moléculas del ADNlc como una fuente de información añadida al cribado prenatal no invasivo y a la biopsia líquida en cáncer. La secuenciación MBD-seq permite no sólo el estudio de la presencia de DMRs, si no también el análisis de los tamaños de fragmentos del ADN metilado. La combinación de ambas se describe por primera vez en esta tesis, mostrando su potencial en la investigación en biomarcadores.

Para estudiar el potencial de dicha tecnología se seleccionó un grupo de estudio que incluía mujeres embarazadas con y sin diagnóstico de sPE. Los síntomas observados en las mujeres que sufren este síndrome son amplios e incluyen hipertensión, proteinuria, trastornos visuales, edema y complicaciones renales y hepáticas. Consideramos que la preeclampsia representa un modelo en el que la que la presencia de la placenta, así como la afectación de

diferentes órganos maternos podría permitir el descubrimiento de cambios epigenéticos asociados a la enfermedad.

Para este fin se analizó el ADNlc metilado en mujeres con embarazados controles y con sPE utilizando secuenciación masiva. El análisis reveló, por primera vez, la existencia de DMRs entre ambos grupos. El hecho de que las DMRs de los dos grupos presenten diferencias de anotación sugiere un potencial papel de la metilación en cambios asociados a la enfermedad. La proporción de islas CpG presentes en las DMRs fue significativamente mayor en las regiones hipermetiladas en sPE, de estas la mayoría correspondían a regiones promotoras del gen. Si bien más del 60% de los promotores de genes humanos presentan islas CpG la mayoría de ellos se encuentran no metilados en circunstancias normales (Baylin and Jones, 2016). Por tanto, la presencia de islas CpG hipermetiladas debe indicar la relevancia de los cambios de metilación en la preeclampsia.

El patrón de metilación observado es similar al descrito en células cancerosas, con una ganancia de metilación en islas CpG en promotores (Robertson, 2005). La placenta comparte este patrón con las células tumorales (Schroeder *et al.*, 2013), y el proceso de placentación ha sido múltiples veces comparado con el cáncer. Si bien el proceso de placentación normal es parecido al cáncer, ya que se produce una invasión exitosa del tejido, en el caso de la preeclampsia ocurre lo contrario. El análisis funcional reveló que los genes hipometilados en sPE se encontraban enriquecidos en rutas relacionadas con el cáncer y la señalización Wnt. Estudios previos han demostrado la metilación de genes supresores de tumores en placentas sanas (Chiu *et al.*, 2007), así como una regulación negativa de la ruta Wnt (Wang *et al.*, 2018) y una sobreexpresión de genes supresores de tumores en placentas de mujeres con preeclampsia comparado con controles (Heikkilä *et al.*, 2005).

Además, los genes con promotores en islas CpG hipermetilados estaban relacionados con varias rutas relacionadas con la patología de la preeclampsia como la ruta de señalización VEGF; cuya inhibición se relaciona con la aparición de los síntomas de preeclampsia (Pandey *et al.*, 2018), o la ruta de contracción del músculo liso vascular; que juega un papel importante en la aparición de la hipertensión en preeclampsia (Goulopoulou and Davidge, 2015).

A parte de términos directamente relacionados con los síntomas de la preeclampsia, el análisis funcional mostró que los genes hipermetilados estaban relacionados con complicaciones a

largo plazo de la misma como diabetes tipo 2 o enfermedades renales y cardiovasculares (Lykke *et al.*, 2009; Jim and Karumanchi, 2017). Los mecanismos de perpetuación de la metilación son fundamentales para el mantenimiento de la identidad celular (Probst *et al.*, 2009). Nuestros datos sugieren que los cambios producidos durante un proceso patológico podrían estar ligados al desarrollo de futuras secuelas.

El análisis funcional también reveló diferencias asociadas a la metilación de TFBS. Específicamente, los TFBS sobre-representados en regiones hipometiladas en sPE correspondían a TFs relacionados con regulación positiva de procesos biológicos, mientras que los que correspondían a TFBS hipermetilados estaban relacionados con regulación negativa. Aunque estas regiones con relevancia en la regulación génica presentaron diferencias en metilación, no podemos inferir el efecto de estos cambios en la expresión génica, ya que se ha demostrado que el efecto de la metilación sobre la unión de los TFs depende del tipo de factor, y puede actuar favoreciendo o evitando su unión, o simplemente sin causar ningún efecto sobre esta (Yin *et al.*, 2017); Xuan Lin *et al.*, 2019).

Podemos concluir que los resultados de análisis funcional de las DMRs correlacionan con cambios previamente observados en preeclampsia. A pesar de que no podemos inferir cambios en la expresión de los genes a partir de nuestros datos, hemos observado que las diferencias en metilación del ADNlc entre embarazos control y con sPE reflejan alteraciones en rutas previamente relacionadas con la enfermedad, así como con sus principales síntomas. Los resultados evidencian el potencial uso del análisis de la metilación del ADNlc en el estudio de la sPE.

Por otro lado, el análisis de los tamaños de los fragmentos de ADNlc reveló diferencias entre ambos grupos, con un enriquecimiento en fragmentos pequeños asociado a sPE. Varios estudios han demostrado que el tejido de origen del ADNlc condiciona el tamaño de las moléculas liberadas a la sangre (Zheng *et al.*, 2012; Sun *et al.*, 2018a; Snyder *et al.*, 2016). En concreto se sabe que el ADNlc origen no hematopoyético (incluyendo la placenta) es de menor tamaño que aquel de las células hematopoyéticas (Zheng *et al.*, 2012). En base a esto, hipotetizamos que el incremento en fragmentos pequeños observados en sPE podría deberse a un incremento en el ADNlc de origen fetal o bien a un aumento en el ADNlc de origen no hematopoyético materno. Tras estudiar la relación entre la contribución fetal y el tamaño de

las moléculas concluimos que las diferencias observadas pudieran ser debidas a un aumento en el ADNlc de origen no hematopoyético, originado en los órganos de la madre en respuesta al síndrome materno de la sPE. Estos datos concuerdan con estudios previos de los que se deduce que la preeclampsia se asocia a un aumento de la concentración fetal de ADNlc (Lo *et al.*, 1999; Martin *et al.*, 2014; Contro *et al.*, 2017), pero éste va acompañado de un aumento mayor en la contribución materna al mismo, causando una reducción en la fracción fetal (Rolnik *et al.*, 2015; Rolnik *et al.*, 2018; Bender *et al.*, 2019; Gerson *et al.*, 2019).

En este trabajo se describe por primera vez la existencia de diferencias en los tamaños de ADNlc metilado entre embarazos control y sPE. Esta identificación pone en evidencia el potencial uso de esta característica del ADNlc en el estudio de la preeclampsia. Aunque más investigación es necesaria para confirmar nuestros hallazgos, estos parecen indicar que el estudio de los patrones de fragmentación del ADNlc puede suponer una fuente de información añadida al estudio de la preeclampsia, proporcionando información de la contribución de los distintos tejidos.

En base a nuestros resultados, que evidencian la utilidad del análisis de la metilación del ADNlc para el estudio de los mecanismos subyacentes a la sPE, exploramos la presencia de diferencias de metilación gen-específicas. Aunque no podemos descartar que algunas de las DMRs se deban a cambios en la contribución diferencial de los tejidos al ADNlc, nuestros resultados muestran cambios en genes previamente relacionados con la preeclampsia. Por ejemplo, entre las regiones hipermetiladas en sPE encontramos algunas en genes relacionados con el control del tono vascular como *ENG* cuya expresión está alterada en preeclampsia (Venkatesha *et al.*, 2006; Sitras *et al.*, 2009) o regiones en genes cuyo potencial en el tratamiento de la preeclampsia ha sido descrito como *VEGFB* (Eddy *et al.*, 2018; Waller *et al.*, 2019). Además, encontramos regiones hipermetiladas en genes como *DNMTA3*, cuya expresión es menor en placentas preeclámpticas (Ma *et al.*, 2018) y cuya hipermetilación se ha relacionado con la preeclampsia por causas inmunológicas (Leavey *et al.*, 2018). O en genes como *MMP23B*, un gen de la familia de las metaloproteasas (MMPs) expresado en la decidua (Anacker *et al.*, 2011). Las MMPs han sido ampliamente estudias en preeclampsia, y se han descrito cambios en la metilación de algunos de sus genes en preeclampsia (Wang *et al.*, 2010;

Rahat *et al.*, 2016a). Estos cambios en el ADNlc parecen indicar un potencial papel de la metilación en la regulación de estos genes importantes en la preeclampsia.

Además, encontramos tres DMRs en el gen *eNOS*; la relevancia de este gen en la preeclampsia se evidencia con el hecho del bloqueo en la producción de oxido nítrico provoca el desarrollo de los síntomas del síndrome materno de la sPE en ratas preñadas (Molnár *et al.*, 1994). Además, la ruta de señalización del oxido nítrico ha sido propuesta como una potencial diana terapéutica para el tratamiento de la preeclampsia mediante el uso de sildenafilo (Paauw *et al.*, 2017). La expresión de este gen se encuentra limitada al endotelio (Förstermann and Sessa, 2012), y la metilación se ha descrito como el mecanismo responsable de esta expresión tejido-específica (Chan *et al.*, 2004; Shirodkar *et al.*, 2013). Entre nuestras DMRs encontramos una región hipermetilada en una isla CpG que, en base a estudios publicados, podría estar relacionada con la regulación de la expresión de este gen (Joo *et al.*, 2013). Teniendo en cuenta las diferencias de metilación observadas, y su posible implicación en la regulación de la expresión del gen, sugerimos que la metilación de *eNOS* representa un marcador interesante para ser estudiado en la preeclampsia.

La preeclampsia grave representa un modelo de patología sistémica, en la que no sólo la salud del feto se ve comprometida, si no que el síndrome que afecta a la madre implica una respuesta inflamatoria generalizada, daño hepático y renal e hipertensión. Por esta razón, consideramos que es el modelo perfecto para evaluar el potencial de la secuenciación MBD-seq del ADNlc en el descubrimiento de cambios epigenéticos específicos de la enfermedad, que podrían proporcionar información del estado global del cuerpo. Trabajos previos han estudiado el ADNlc en esta patología, pero estos han estado centrados en el análisis de las diferencias de concentración del ADNlc. En este trabajo, vamos un paso más allá y exploramos la hipótesis de que el ADNlc de mujeres con preeclampsia debe llevar consigo información con relevancia para la patología. Hasta donde sabemos, este estudio es el primero en reportar la existencia de diferencias en la metilación del ADNlc entre embarazos sanos y con preeclampsia, así como en los tamaños de las moléculas de ADNlc metilado. Los resultados de esta tesis refuerzan dicha hipótesis y sientan las bases para el estudio de la metilación del ADNlc, así como de otras de sus características, en esta y otras enfermedades en las que múltiples órganos se ven afectados.

CONCLUSIONES:

- La secuenciación del ADNlc metilado reveló la presencia de regiones diferencialmente metiladas entre embarazos control y embarazos con preeclampsia grave.

- El enriquecimiento en promotores e islas CpG observado en las regiones hipermetiladas respalda el papel de la metilación del ADN en los mecanismos subyacentes a la preeclampsia grave.

- Las diferencias de metilación en ADNlc entre preeclampsia grave y embarazos control reflejan la alteración en rutas previamente relacionadas con la enfermedad como rutas en cáncer, la ruta de señalización Wnt, la ruta de señalización del VEGF y la ruta de contracción del músculo liso vascular.

- El análisis de asociación gen-enfermedad reveló un enriquecimiento en genes relacionados con los síntomas y secuelas de la preeclampsia entre las regiones hipermetiladas. Estos datos sugieren una implicación de la metilación en los mecanismos que predisponen a las mujeres con preeclampsia a desarrollar enfermedades cardiovasculares y metabólicas con posterioridad.

- El análisis de los tamaños de fragmentos mostró diferencias significativas entre embarazos control y con preeclampsia grave, con un enriquecimiento en fragmentos pequeños en la preeclampsia negativamente correlacionado con la fracción fetal. Estos datos sugieren un incremento en la contribución materna al ADNlc en respuesta a la enfermedad.

- La secuenciación del ADNlc enriquecido mediante el uso de proteínas de unión a metilcitosina permite la identificación de diferencias de metilación en una rutina de diagnóstico clínico.

- El estudio de características epigenéticas en el ADNlc representa una fuente de información explotable en enfermedades que implican la afectación de múltiples órganos como la preeclampsia grave. En concreto, el uso de diferencias epigenéticas amplia la

aplicación de la biopsia líquida a la detección de marcadores en enfermedades sin elevadas tasas de mutación.

INDEX

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACOG	American College of Obstetricians and Gynecologists
cfDNA	cell-free DNA
DMPs	Differentially Methylated Probes
DMRs	Differentially Methylated Regions
DNMTs	DNA methyltransferases
ENCODE	Encyclopedia of DNA elements
eNOS	endothelial Nitric Oxide Synthase
EOPE	Early-Onset Preeclampsia
FDR	False Discovery Rate
FGR	Fetal Growth Restriction
FLK1	Protein-Tyrosine Kinase Receptor Flk-1
FLT1	fms-like tyrosine kinase
GAD	Gene Association Database
HIF1α	Hypoxia Indicuble Factor 1α
HLA-C	Human Leukocyte Antigen Type C
IUGR	Intrauterine Growth Restriction
KEGG	Kyoto Encyclopedia of Genes and Genomes
KIRs	Killer-cell immunoglobulin-like receptors
LOPE	Late-Onset Preeclampsia
MBD	Methyl-Binding Domain
MBD-seq	MBD enriched DNA sequencing
MeCP2	Methyl-CpG-binding Protein 2
MeDIP-seq	Methylated immunoprecipitation sequencing
methyl-CpG	Methylcitosine
MHC	Major Histocompatibility Complex
MMPs	Matrix Metalloproteases
MRE-seq	Methylation sensitive Restriction Enzyme sequencing
mtDNA	Mitochondrial DNA

NGS	Next Generation Sequencing
NIPT	Non-invasive Prenatal Testing
NK	Natural Killers
ORA	Overrepresentation enrichment Analysis
PAPP-A	Pregnancy-Associated Plasma Protein-A
PBC	Peripheral Blood Cells
PCR	Polymerase Chain Reaction
PE	Preeclampsia
PGCs	Primordial Germ Cells
PlGF	Placental Growth Factor
RRBS	Reduced Representation Bisulfite Sequencing
sENG	soluble Endoglin
sFLT1	soluble Fms-like tyrosine kinase 1
sPE	Severe Preeclampsia
TET	ten-eleven translocation methylcytosinedioxygenase family enzymes
TF	Transcription Factor
TFBS	Transcription Factor Binding Sites
TGFβ	Transforming Growth Factor Beta
TGFβ1	Transforming Growth Factor Beta 1
TGFβ3	Transforming Growth Factor Beta 3
TSGs	Tumor Suppressor Genes
UB-WBC	Umbilical Blood White Cells
UVEC	umbilical vein endothelial cells
VEGF	Vascular Endothelial Growth Factor
VEGFR1	Vascular Endothelial Growth Factor Receptor 1
WGBS	Whole Genome Bisulfite Sequencing
ZBTB33	Kaiso or Zinc Finger and BTB Domain Containing 33

I. INTRODUCTION

I. INTRODUCTION

1. PREECLAMPSIA

1.1 CLASSIFICATION OF HYPERTENSIVE DISORDERS IN PREGNANCY

Hypertension disorders of pregnancy complicate up to 10% of pregnancies worldwide (ACOG, 2013). In 1972, the Task Force on Hypertension in Pregnancy of the American College of Obstetricians and Gynecologists (ACOG) introduced a classification for hypertensive disorders of pregnancy, later updated in 2013. This classification divides hypertensive disorders of pregnancy into four categories: chronic hypertension, gestational hypertension, preeclampsia-eclampsia and chronic hypertension with superimposed preeclampsia (ACOG, 2013).

Chronic hypertension is defined by the presence of high blood pressure preceding the conception or detected before the 20th week of pregnancy. **Gestational hypertension** is characterized by the new-onset of high blood pressure after 20 weeks of pregnancy in the absence of proteinuria that disappears postpartum**. Preeclampsia** is the most common of the pregnancy-specific hypertensive disorders (ACOG, 2019). It is classically defined by new -onset hypertension and new-onset proteinuria, occurring after 20 weeks of gestation. However, preeclampsia can be diagnosed in the absence of proteinuria in some women who present multisystemic signs (ACOG, 2013). Development of grand mal seizures in a woman with preeclampsia is defined as **eclampsia** and is considered an end stage of the disease (Phipps *et al.*, 2019). **Chronic hypertension with superimposed preeclampsia** refers to the development of preeclampsia in women with chronic hypertension (ACOG, 2013).

1.2 PREECLAMPSIA: DEFINITION AND DIAGNOSIS

Preeclampsia is a hypertensive syndrome that complicates 2-8% of pregnancies globally (Duley, 2009). Along with eclampsia, it is one of the top 3 leading causes of maternal morbidity and mortality worldwide (Ghulmiyyah and Sibai, 2012). Preeclampsia can occur in molar pregnancies, when a fetus is absent, showing that placenta, but not fetus, is required for the development of the disease (Acosta-Sison, 1956). Thus, placenta is considered essential for the development and remission of the syndrome. In fact, the delivery of the placenta is the only definitive treatment for preeclampsia (Chaiworapongsa *et al.*, 2014).

In the ACOG classification of 1972, preeclampsia was initially diagnosed by the new onset of hypertension and proteinuria after 20 weeks of pregnancy in a woman with previously normal blood pressure. Hypertension is defined as systolic blood pressure ≥ 140 mmHg, diastolic blood pressure ≥ 90 mmHg, or both. The 2013 update in the classification eliminated proteinuria as a requirement for diagnosis. In the absence of proteinuria, diagnosis requires hypertension together with the presence of any of the following: new-onset thrombocytopenia, renal insufficiency, pulmonary edema, impaired liver function or visual and cerebral disturbances (ACOG, 2013) (Table 1).

Table 1: Preeclampsia diagnostic criteria (from ACOG, 2013)

Blood pressure	• Greater than or equal to 140 mm Hg systolic or greater than or equal to 90 mm Hg diastolic on two occasions at least 4 hours apart after 20 weeks of gestation in a woman with a previously normal blood pressure • Greater than or equal to 160 mm Hg systolic or greater than or equal to 110 mm Hg diastolic, hypertension can be confirmed within a short interval (minutes) to facilitate timely antihypertensive therapy
and	
Proteinuria	• Greater than or equal to 300 mg per 24-hour urine collection (or this amount extrapolated from a timed collection) or • Protein/creatinine ratio greater than or equal to 0.3* • Dipstick reading of 1+ (used only if other quantitative methods not available)
Or in the absence of proteinuria, new-onset hypertension with the new onset of any of the following:	
Thrombocytopenia	• Platelet count less than 100,000/microliter
Renal insufficiency	• Serum creatinine concentrations greater than 1.1 mg/dL or a doubling of the serum creatinine concentration in the absence of other renal disease
Impaired liver function	• Elevated blood concentrations of liver transaminases to twice normal concentration
Pulmonary edema	
Cerebral or visual symptoms	

* Each measured as mg/dL.

Preeclampsia with severe features is defined as preeclampsia with any of the following features: systolic blood pressure ≥160 mmHg or diastolic blood pressure ≥110 mmHg; thrombocytopenia (platelet count <100,000/microliter); impaired liver function (abnormally elevated liver enzymes to twice the normal concentration or severe persistent right upper quadrant or epigastric pain unresponsive to medication and not accounted for by alternative medication); progressive renal failure (serum creatinine concentration level of >1.1 mg/dL or doubling of the serum creatinine concentration in the absence of other renal disease), pulmonary edema or new-onset cerebral or visual disturbances (ACOG, 2013).

1.3 PATHOGENESIS

Preeclampsia pathogenesis involves two stages (Figure 1): abnormal placentation, and development of maternal syndrome. Indeed, poor placentation results in ischemia of the organ, which releases factors in maternal circulation that cause systemic endothelial dysfunction and maternal signs of the condition (Phipps *et al.*, 2019; Chaiworapongsa *et al.*, 2014; Redman and Sargent, 2005).

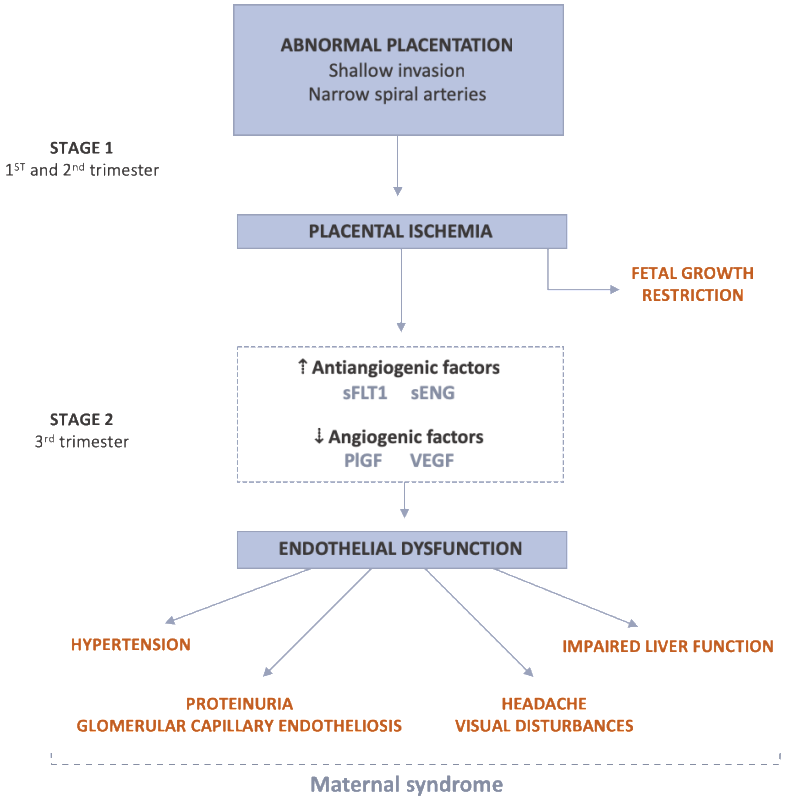

Figure 1. Two-stage model for preeclampsia pathogenesis. During the 1st stage, failure of physiological spiral artery transformation causes an abnormal placentation. As a result, placental perfusion is reduced, causing fetal growth restriction. During the 2nd stage, maternal syndrome is produced by an imbalance in angiogenic/antiangiogenic factors released by the ischemic placenta to maternal circulation. It is characterized by systemic inflammatory response, endothelial dysfunction, hypertension, proteinuria, visual disturbances, edema, liver and kidney failure. *sFLT1: Soluble fms-like tyrosine kinase 1, sENG: Soluble endoglin, PlGF: Placental growth factor, VEGF: Vascular endothelial growth factor*

1.3.1 Placentation

During the first stages of implantation nutrition is provided by secretions of the uterine glands under hypoxic conditions, known as histiotrophic nutrition (Burton *et al.*, 2002). Is not until 10-12 weeks of pregnancy when uteroplacental circulation is fully established and there is a shift into hemotrophic nutrition (Jauniaux *et al.*, 2000). In the early stages, low oxygen tension

promotes trophoblast proliferation and avoids differentiation to an invasive phenotype by mechanisms involving hypoxia inducible factor 1α (HIF1α) and transforming growth factor beta 3 (TGFβ3). During the 1st trimester, extravillous trophoblasts start to invade maternal tissues and can be classified as interstitial trophoblasts and endovascular trophoblasts. Interstitial trophoblasts invade the decidua and adjacent myometrium by secreting proteolytic enzymes that degrade extracellular matrix and activating proteinases already present in the endometrium. In addition, these cells aggregate around spiral arteries, possibly as a preparation for endovascular invasion. Endovascular trophoblasts first enter the spiral arterial lumen and plug the arteries until week 9, limiting utero-placental circulation (Cunningham *et al.*, 2014).

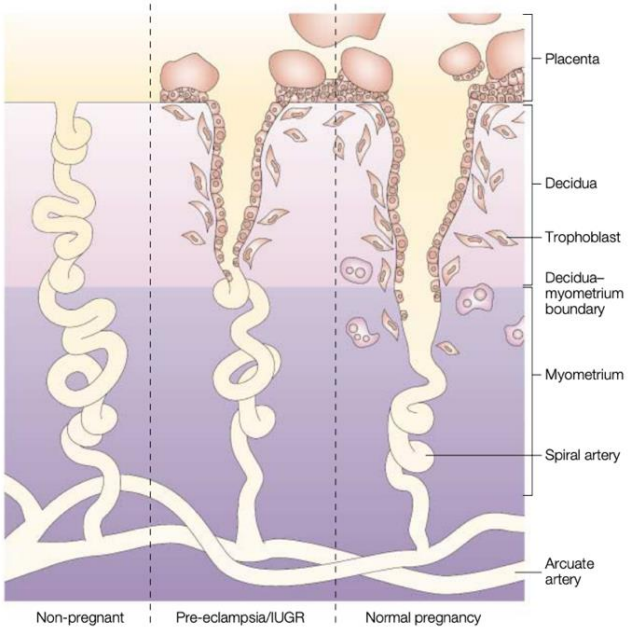

Figure 2. Abnormal placentation and preeclampsia. During normal pregnancy, invasive trophoblasts migrate, destroy vascular endothelium and transform spiral arteries, widening their diameter. This invasion takes place in 2 waves: the first one affects spiral arteries up to the border between decidua and myometrium, while the 2nd one reaches intramyometrial segments of spiral arteries. In preeclampsia, only superficial decidua segments are transformed. Incomplete spiral artery transformation results in an ischemic placenta. *IUGR: Intrauterine growth restriction.* Figure from Moffett-King, 2002; reproduced with the permission from Springer Nature.

Maintenance of hypoxia is critical because the fetus, which is undergoing organogenesis, is particularly vulnerable to free radicals (Redman and Sargent, 2005). At the end of the first trimester these plugs are progressively dislocated. The flow of maternal blood progressively leads to hemotrophic nutrition that will not be complete until intramyometrial segments of spiral arteries (Cunningham *et al.*, 2014) (Figure 2). In preeclampsia, spiral artery transformation is incomplete; only superficial decidua segments are transformed, while

arterial myometrial segments remain narrow, resulting in an ischemic placenta (Young *et al.*, 2010) (Figure 2).

Aberrant regulation of trophoblast differentiation mediated by hypoxia might explain the failure of trophoblasts to acquire the invasive phenotype. During normal pregnancy, hypoxia regulates trophoblast differentiation by mechanisms involving HIF1α and TGFβ3. Both are highly expressed during early placentation, when oxygen tension is low, and their expression is reduced at the end of the first trimester (10–12 weeks) when oxygen tension increases (Caniggia *et al.*, 2000). In preeclamptic placentas expression of HIF2α, HIF1α and TGFβ3 protein is increased (Caniggia *et al.*, 1999; Rajakumar *et al.*, 2004). Studies in villous explants indicate that a normal invasive phenotype is restored in tissue from patients with preeclampsia by inhibition of trophoblast TGFβ3 expression (Caniggia *et al.*, 1999). In addition, antisense inhibition of HIF-1α downregulates TGFβ3 and triggered trophoblast invasive phenotype in first trimester placentas (Caniggia *et al.*, 2000).

Failure in maternal-fetal immune recognition could also explain defective placentation in preeclampsia. Fetal implantation is a unique scenario, in which the development of a semiallogenic entity, the fetus, is permitted by the maternal immune system (Young *et al.*, 2010). Preeclampsia has been considered as a form of maternal immune rejection of the genetically unrecognized fetus. Immune maladaptation contributes to an inadequate invasion of the trophoblasts into uterine decidua (Young *et al.*, 2010). This hypothesis explains why preeclampsia occurs typically in the first pregnancy (Chaiworapongsa *et al.*, 2014).

Understanding the immune recognition process that takes place during placentation necessitates knowledge of the key players. In the placenta, the cells that are in contact with maternal blood (villous trophoblasts) lack major histocompatibility complex (MHC) class I and II molecules, while extravillous trophoblasts, the ones that invade uterus, express MHC class I molecules (Moffett-King, 2002). The predominant immune population on the maternal side comprises the natural killer cells (NK), instead of T and B cells, typical of the adaptive immune response (Moffett-King, 2002). More specifically, the process involves human leukocyte antigen type C (HLA-C) molecules and their receptors on uterine NK cells, killer-cell immunoglobulin-like receptors (KIRs). HLA-C molecules bind KIR receptors making uterine NK

cells release factors that promote trophoblast invasion. This binding is stronger with HLA-C 1 than with HLA-C 2. Moreover, secretion of these factors increases when HLA-C antigens are bound to haplotype B KIRs (stimulatory) and reduced when they bind to haplotype A KIRs (Redman and Sargent, 2010). Maternal KIR haplotype BB increases the possibilities of an adequate placentation. Accordingly, KIR haplotype AA mothers combined with HLA-C2 fetuses present higher risk of preeclampsia (Hiby *et al.*, 2010; Hiby *et al.*, 2004).

1.3.2 Maternal syndrome

The second stage consists of the consequences of poor placentation: reduced placental perfusion leads to fetal growth restriction. Simultaneously, on the maternal part, the ischemic placenta releases antiangiogenic factors and other inflammatory mediators into maternal circulation. This release causes a systemic inflammatory response, endothelial dysfunction and the maternal signs of the condition: hypertension, proteinuria, visual disturbances, edema, liver and kidney failure (Redman and Sargent, 2005; Young *et al.*, 2010) (Table 1, Figure 1).

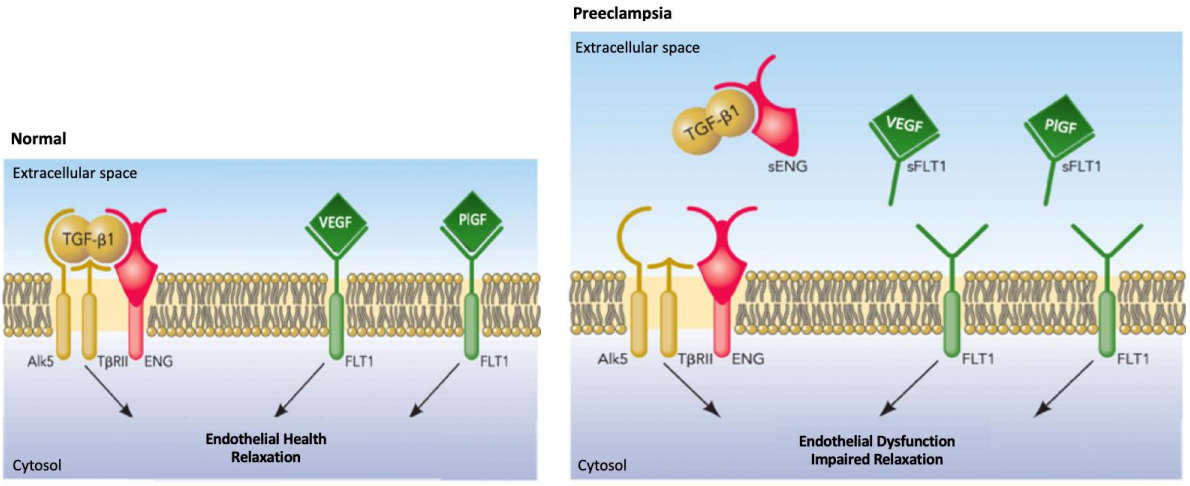

Figure 3. Role of antiangiogenic factors in preeclampsia. During normal pregnancy, vascular homeostasis is maintained by a balanced between angiogenic and antiangiogenic factors. In preeclampsia, placental release of antiangiogenic factors is increased. Excess of sENG binds TGFβ and antagonizes its effect on endothelial cells. In the same way, sFLT1 present in maternal circulation binds to pro-angiogenic molecules (VEGF and PlGF), impairing the angiogenic effect. As a result, VEGF and TGFβ1 downstream signaling is inhibited causing endothelial cell dysfunction and impairing relaxation. *TGF*β1*: Transforming growth factor beta 1, Alk5: Transforming growth factor beta Receptor 1, T*βRII: *Transforming growth factor beta receptor 2, ENG: Endoglin, FLT1: Fms-like tyrosine kinase 1, VEGF: Vascular endothelial growth factor, PlGF: Placental growth factor, sENG: Soluble endoglin, sFLT1: Soluble fms-like tyrosine kinase 1.* Figure adapted from Wang *et al.,* 2009.

Imbalance in angiogenic factors plays an important role in preeclampsia (Figure 3). Antiangiogenic factor soluble fms-like tyrosine kinase 1 (sFLT1) is produced in placenta and is released into maternal circulation. It acts as a ligand trap for pro-angiogenic molecules vascular endothelial growth factor (VEGF) and placental growth factor (PlGF), preventing their binding to VEGF receptor 1 (VEGFR1) and the subsequent angiogenic effect. In preeclampsia, sFLT1 is upregulated in placenta, and higher levels of circulating sFLT1 associated with decreased levels of circulating VEGF and PlGF are found in maternal blood of preeclamptic pregnancies (Maynard *et al.*, 2003). In rodents, sFLT1 overexpression results in the development of maternal preeclampsia symptoms, including hypertension, proteinuria and glomerular capillary endotheliosis (Maynard *et al.*, 2003).

Soluble endoglin (sENG) is an antiangiogenic protein that is also associated with preeclampsia (Figure 3). sENG of placental origin is found in blood of pregnant women and is increased in preeclamptic pregnancies, correlating with the severity of the disease (Venkatesha *et al.*, 2006). It is detected at high concentrations in endothelial cells and syncytiotrophoblasts and regulates TGFβ1 and TGFβ3 signaling. TGFβ induces vasorelaxation through activation of endothelial nitric oxide synthase (eNOS); sENG interferes with receptor binding and downstream signaling in endothelial cells and eNOS activation. Moreover, sEng cooperates with sFLT1 to induce endothelial dysfunction. In pregnant rats, adenoviral expression of sENG causes a significant change in mean arterial pressure and modest proteinuria compared with the bigger increase in arterial pressure and proteinuria obtained with sFLT1 expression. Compared to the expression of both independently, expression of sENG + sFLT1 leads to higher levels of arterial pressure and proteinuria, and to a decrease in platelet count, impairment of liver function, vascular damage of the placenta and fetal growth restriction (Venkatesha *et al.*, 2006). Since the alterations in the concentrations of sFlt-1, PlGF, and sENG in maternal circulation precede the clinical onset of preeclampsia, they are used, in combination with other measurements such as serum pregnancy-associated plasma protein-A (PAPP-A), uterine pulsatility index and mean arterial pressure, as predictors of the disease (ACOG, 2013).

As research efforts continue toward defining the pathogenesis of preeclampsia and identifying avenues for averting this condition, epigenetic studies may provide a key opportunity for discovery. Since DNA methylation have been proved to have a role during different disease

processes (Robertson, 2005; Jin and Liu, 2018; Fernandez *et al.*, 2012), its analysis would represent a new approach in the study of preeclampsia pathogenesis.

2. EPIGENETICS

Epigenetics is the study of heritable (and often modifiable) information beyond that encoded in the DNA sequence. It encompasses chromatin and DNA modifications, as well as other mechanisms of transcriptional regulation that act in the context of chromatin (Greally, 2018). Epigenetic features, including DNA methylation, histone modifications, promoter–enhancer interactions, and noncoding RNAs, regulate genome function distinctly in different cell types.

2.1 DNA METHYLATION

DNA methylation is well conserved among most plant, animal and fungal systems, and is the most studied epigenetic modification (Smith and Meissner, 2013). This covalent modification consists of the addition of a methyl (CH3) group on cytosine nucleotides. DNA methylation is a dynamic process: methylation marks can be synthesized de novo, maintained or removed. These processes are mediated by DNA methyltransferases (DNMTs) and DNA demethylases. DNMT3A and DNMT3B write the methylation pattern on the DNA through de-novo methylation, DNMT1 controls the maintenance of methylation following DNA replication, and ten-eleven translocation methylcytosinedioxygenase family enzymes (TET) remove methyl groups (Chen *et al.*, 2017).

DNA methylation exclusively occurs on cytosine, and almost always on cytosines that precede a guanine nucleotide, known as CpG sites. Mammalian genomes are considered as globally CpG-depleted. The human genome contains 28 million CpGs, which represents less than 1% of the genome. Of them, 60–80% are typically methylated, except for CpG islands that are largely resistant to methylation (Smith and Meissner, 2013). Thus, the genome-wide DNA methylation landscape is bimodal, with the large majority of CpG methylated and CpG islands unmethylated (Bergman and Cedar, 2013). CpG islands are regions of the genome that exhibit strong enrichment for CpG dinucleotides. The human genome contains around 30,000 CpG islands, which range in length from several hundred to several thousand base pairs. CpG islands associated with promoters nearly always remain unmethylated, while many of the CpG

islands lying within gene bodies become methylated during development and differentiation (Jeziorska et al., 2017).

2.2 FUNCTION OF DNA METHYLATION

2.2.1 Embryo development

The embryonic development landscape is characterized by dynamic methylation changes, in contrast to the stability of the methylome in somatic cells (Luo *et al.*, 2018). Unlike DNA sequence, which is directly inherited from the parents, a new methylation profile is established in each embryo (Dor and Cedar, 2018) (Figure 4). This process starts with the erasing of DNA methylation patterns derived from the parents, which takes place after fertilization. Around the time of embryo implantation, methylation begins resetting. The generation of the new profile involves two stages. First, a dramatic wave of de novo methylation occurs in the entire embryonic genome. Surprisingly, CpG islands escape from this process and remain unmethylated, which generates a bimodal methylation pattern (Cedar and Bergman, 2012). This pattern is maintained in every cell (Dor and Cedar, 2018). The second stage of resetting genome methylation involves additional methylation changes that occur in a tissue-specific or gene-specific nature during cell-lineage differentiation and organogenesis (Dor and Cedar, 2018). Discrete changes take place during the post-implantation stage: de novo methylation occurs in some genes that were unmethylated and are turned off during development in a particular cell type, while demethylation occurs in other regions during differentiation events.

This newly established pattern is maintained for the rest of the organism's life, generating a stable tissue-specific chromatin template. In spite the stability of this pattern, DNA methylation is modifiable, in fact, DNA methylation changes occur across the lifespan by external influences from the environment, such as diet, trauma, exercise, or disease (Dor and Cedar, 2018). Imprinted genes are an exception to these events; imprinted genes are expressed only from one of the two inherited chromosomes, as determined by the parent of inheritance (Moore *et al.*, 2013). These genes escape the genome-wide reprogramming that occurs after fertilization, failing to undergo either methylation pattern erasure in the preimplantation stage, or de novo methylation at the time of implantation (Cedar and Bergman, 2012; Barlow and Bartolomei, 2014).

2.2.2 Gametogenesis

During gametogenesis, methylation plays an important role in the epigenetic reprogramming of the germline, a fundamental process in mammals that appears highly conserved (von Meyenn and Reik, 2015). At the post-implantation stage, primordial germ cells (PGCs), the precursors of sperm and oocytes, migrate to the developing gonads (Figure 4). At that moment PGCs have the bimodal methylation pattern acquired after demethylation and methylation waves. During migration a new demethylation event occurs, in which almost all DNA methylation is lost from CpG islands, transcription start sites, gene bodies and surrounding intergenic regions (von Meyenn and Reik, 2015). This epigenetic resetting only includes erasure of imprinted genes in the PGCs; the rest of the somatic cells maintain the inherited imprinting on the same parental chromosome (Barlow and Bartolomei, 2014). Following demethylation and differentiation of the PGCs, methylation is imposed on the imprinted regions in a sex-specific manner at late stages of gametogenesis, either in the oocytes or sperm (Bartolomei and Ferguson-Smith, 2011), that will be maintained in the into adulthood of the offspring.

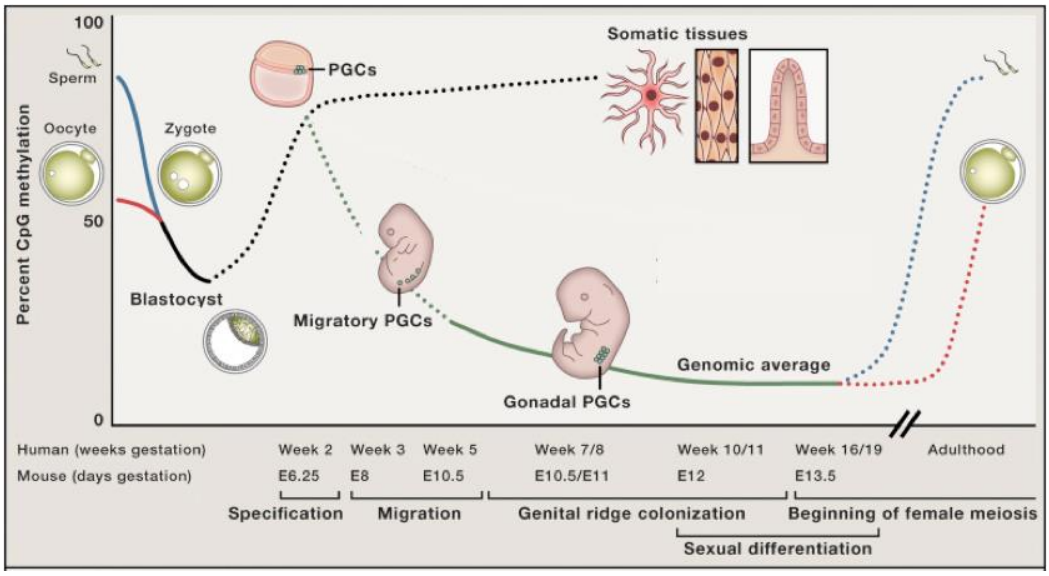

Figure 4. Methylation during embryo development and gametogenesis. After fertilization, inherited methylation patterns are erased via global demethylation that takes place in both paternal (blue) and maternal (red) genomes. Around the blastocyst stage, de novo methylation starts in a process involving two stages. First, the entire embryonic genome undergoes a dramatic wave of methylation, during which CpG islands remain non-methylated. During the second stage, methylation changes occur in a tissue- or gene-specific manner, establishing a new pattern that is maintained for the rest of the organism's life and thereby generating a stable tissue identity. Additionally, during the post-implantation stage, primordial germ cells (PGCs) migrate to the gonads while a new demethylation wave takes place (green dotted line). After that, at the late stages of gametogenesis, a new methylation pattern is generated in a sex-specific manner (red and blue dotted lines). Figure adapted from von Meyenn and Reik, 2015

2.2.3 Gene expression

Although traditionally methylation has been considered as a gene silencing mechanism, the role of methylation in gene expression control is now known to be dependent on its position on the transcriptional unit (Jones, 2012). DNA methylation within intergenic regions represses gene expression, probably as a mechanism to maintain genome stability and avoid harmful expression of repetitive and transposable elements present in these regions (Moore *et al.*, 2013). In the context of genes, the effect of **gene body** methylation remains unclear. While some studies have reported an association between gene body methylation and gene silencing (Moore *et al.*, 2013), studies in cancer have shown a positive correlation between methylation and expression (Yang *et al.*, 2014; Arechederra *et al.*, 2018). This is supported by the fact that active chromosome X displays more than two times the allele-specific methylation than inactive X, and this methylation is concentrated at gene bodies (Hellman and Chess, 2007). Within **promoter regions,** it is necessary to differentiate between CpG islands promoters and non-CpG islands promoters. CpG islands are typically unmethylated, even when gene transcription is not active (Deaton and Bird, 2011). That genes with CpG islands regions are associated with fewer nucleosomes (Fenouil *et al.*, 2012) and that many transcription factors bind to GC-rich regions (Zhu *et al.*, 2016) suggest that CpG islands promoters are regulated by transcription factors binding and other mechanisms involving chromatin accessibility. In contrast, CpG islands promoter methylation is associated with genes with long-term repressed states, where methylation results in stable silencing of gene expression (Jones, 2012; Mohn *et al.*, 2008). CpG islands methylation is not the initiating event in gene silencing, but it is instrumental in the maintenance of the silent state by blocking the expression (Deaton and Bird, 2011). On the other hand, methylation levels of non-CpG islands promoters are more dynamic, and represent an important mechanism of regulation in these promoters, leading to gene silencing (Han *et al.*, 2011; Hartung *et al.*, 2012; Øster *et al.*, 2013).

Methylation may reduce gene expression itself, but there are some proteins that have high affinity for methylcytosine (methyl-CpG) and influence gene expression. These proteins recognize and bind to methyl-CpG acting as readers of methylation and leading to gene silencing (Moore *et al.*, 2013; Zhu *et al.*, 2016). The methyl-binding domain (MBD) protein family comprises five mammalian members: methyl-CpG-binding protein 2 (MeCP2) and MBD1-4. All but MBD3 bind to methylated DNA in a non-sequence-specific manner (Zhu *et al.*,

2016). MBD1, MBD2 and MeCP2 are associated with gene silencing; they contain a transcriptional repressor domain to which repressor complexes are bound (Moore *et al.*, 2013). The classical view of methylated DNA-protein interactions was that only proteins with MBDs can recognize and bind to methylated CpGs. However, transcription factors lacking these domains can interact with methylated DNA in a sequence-specific manner (Zhu *et al.*, 2016). Zinc-finger proteins Kaiso (ZBTB33), ZBTB4 and ZBTB38 are transcription regulators that bind to methylated CpG and repress transcription in a methylation-dependent manner (Moore *et al.*, 2013).

2.2.4 Methylation in disease

DNA methylation is crucial in regulating many cellular processes from embryonic development, when a new methylation pattern is established in each embryo, to adulthood when it maintains this pattern, regulates gene transcription and maintains genome stability. Thus, a correctly established and maintained methylation pattern seems to be fundamental for the proper development and functioning of an organism. Increasing evidence of the involvement of DNA methylation in the development and progression of human diseases underscores its importance (Robertson, 2005; Jin and Liu, 2018).

The role of DNA methylation in human diseases was first explored in the context of genomic imprinting (Jin and Liu, 2018). Imprinted genes are expressed from a single allele. Differential allele-specific DNA methylation occurs in imprinted regions as the result of de novo methylation in one gamete that is maintained into adulthood of the offspring. These regions can gain or lose their DNA methylation marks in a process called loss of imprinting, and it is implicated in many imprinted disorders such as Beckwith-Wiedemann syndrome, Prader-Willi syndrome and Angelman syndrome (Jin and Liu, 2018; Bartolomei and Ferguson-Smith, 2011).

Although methylation changes are implicated in non-cancerous diseases such as Alzheimer's, dementia with Lewy bodies, atherosclerosis, myopathies and autoimmune disorders (Fernandez *et al.*, 2012), the most studied case of methylation and disease is cancer.

Tumor cells share an abnormal methylation pattern, with demethylation of many non–CpG islands regions and some CpG islands undergoing de novo methylation (Robertson, 2005; Baylin and Jones, 2016). The fact that all tumors show similar changes in DNA methylation

suggests that this may be a basic element in cancer biology and tumor pathology (Bergman and Cedar, 2013).

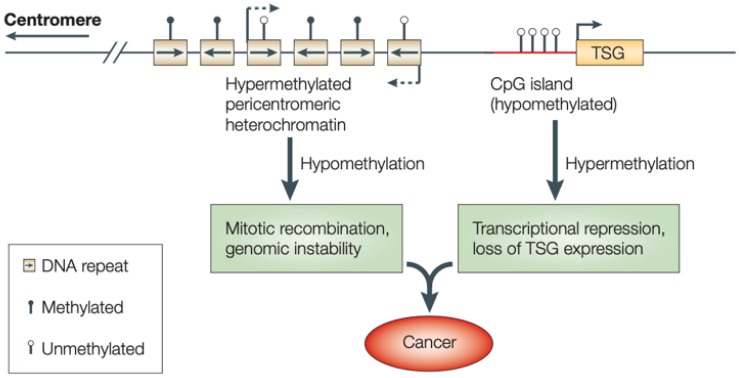

Figure 5. Methylation changes in cancer. Diagram shows a representative part of genomic DNA from normal cells. A hypermethylated region corresponds to repeat-rich pericentromeric heterochromatin; a hypomethylated region corresponds to CpG island of an actively transcribed tumor-suppressor gene (TSG). During tumorigenesis, global demethylation occurs, affecting repeat-rich heterochromatin regions that predispose cells to genomic instability. Moreover, de novo methylation of CpG islands, typically occurring at TSGs, leads to transcriptional silencing. Figure from Robertson, 2005; reproduced with permission from Springer Nature.

Loss of genomic methylation is an early event in tumorigenesis that takes place in repetitive elements of the genome (Figure 5). When becoming hypomethylated, these repetitive regions predispose cells to genomic instability, a hallmark of tumor cells. Moreover, global DNA hypomethylation is associated with poorer prognosis in cancer patients (Li *et al.*, 2014). Apart from this global demethylation, gene-specific hypomethylation events also occur in cancer. Global demethylation may predispose cells to genomic instability, while gene-specific changes that occur later may allow tumor cells to adapt to their environment and promote metastasis (Robertson, 2005; Baylin and Jones, 2016). Simultaneously with demethylation events, de novo methylation events occur in cancer cells. Aberrant hypermethylation mostly occurs in CpG islands and causes gene silencing in a targeted manner (Baylin and Jones, 2016). CpG islands in genes involved in cell-cycle regulation, tumor cell invasion, DNA repair, chromatin remodeling, cell signaling, transcription and apoptosis are hypermethylated and silenced in nearly every tumor type. Together these changes promote cell growth, genetic instability and metastasis (Robertson, 2005). Remarkably, methylation changes that occur during tumorigenesis are the same that occur during aging in normal cells. This process takes place in all cells of the body, but with different rates for each tissue. Indeed, the risk of cancer depends on the degree of abnormal age-related methylation changes (Dor and Cedar, 2018).

In the CpG island context, CpG islands hypomethylation-mediated activation of oncogenes has also been documented in several cancer types (Ferreira and Esteller, 2018).

2.3 DNA METHYLATION STUDY

Given its roles in both normal physiology and disease pathogenesis, the study of DNA methylation can provide important insights to myriad processes during health and disease. DNA methylation analysis can be generally divided into targeted, when methylation status of some genes or regions of interest is assessed, typically by classical molecular or biochemical techniques, or genome-wide, when global DNA methylation profiling is performed by high-throughput microarray or next generation sequencing (NGS) (Yong *et al.*, 2016).

Research on methylation in humans typically focuses on the identification of differentially-methylated regions (DMRs). These regions are defined as genomic loci that present different methylation status between compared samples. DMRs have been identified in multiple scenarios such as between different healthy tissues (Fernandez *et al.*, 2012; Kundaje *et al.*, 2015), young and older people (Dor and Cedar, 2018), and of course between health and disease states (Fernandez *et al.*, 2012). The identification of DMRs among multiple tissues provides a comprehensive survey of epigenetic differences in human tissues. The discovery of DMRs in disease, on the other hand, can provide information on pathogenesis and potentially identify biomarkers of disease (Neidhart, 2015).

2.3.1 Methylation study methods

Different approaches are available for profiling DNA methylation. In this section the most used are explained in detail and are summarized in Figure 6.

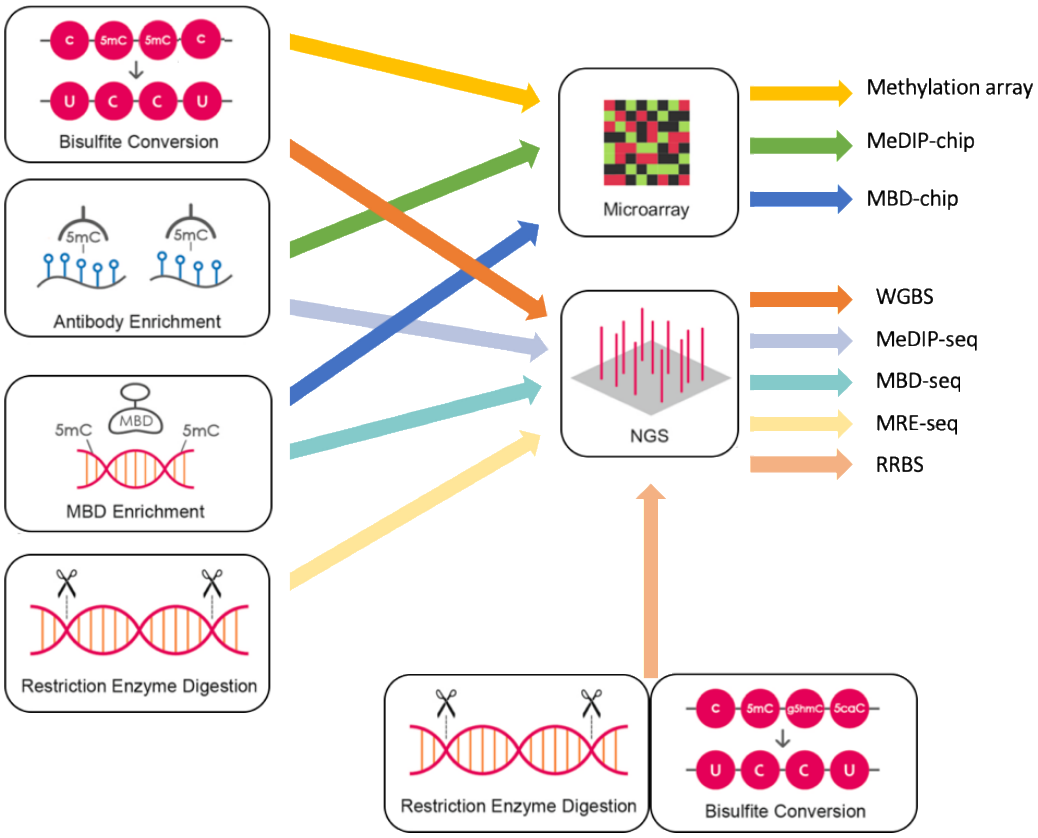

Figure 6. Summary of methylation study methods. For methylation analysis DNA can be converted by bisulfite treatment, enriched using methyl-binding domain (MBD) enrichment or methylated DNA immunoprecipitation (MeDIP), digested using methylation sensitive restriction enzymes (MREs) or digested followed by bisulfite conversion. After that, the resulting product is analyzed by microarray or next generation sequencing (NGS). Whole-genome bisulfite sequencing (WGBS) consists of sequencing of bisulfite-treated DNA. Methylated DNA enriched either by antibodies or by MBD enrichment can be analyzed in microarray platforms (MeDIP-chip or MBD-chip) or NGS (MeDIP-seq or MBD-seq). DNA digested with MREs and directly analyzed by sequencing is called MRE-seq, while when digestion is followed by bisulfite conversion is called reduced representation bisulfite sequencing (RRBS). *5mC: 5-methylcytosine.* Figure adapted from Yong *et al.*, 2016.

2.3.1.1 Bisulfite-based methods

Bisulfite treatment converts unmethylated cytosines to uracil by deamination, while methylated cytosines remain unchanged (Figure 7), allowing the study of methylation at single-base resolution. Bisulfite-treated DNA is used to study the "methylome" by hybridization in methylation arrays, whole-genome bisulfite sequencing (WGBS) or reduced-representation bisulfite sequencing (RRBS).

Bisulfite conversion of DNA causes significant DNA degradation through a chemical reaction that takes place under harsh conditions such as low pH, high temperature and elongated incubation times (*DNA Methylation Protocols*, 2018). During bisulfite treatment two types of

conversion error can occur: inappropriate conversion and failed conversion. Failed conversion occurs when an unmethylated cytosine fails to be converted to uracil, resulting in false-positive methylation signals. On the other hand, inappropriate conversion occurs when a methylated cytosine is converted into a thymine, causing false-positive signals (Genereux *et al.*, 2008). Controlling reaction conditions (temperature, molarity, incubation times) is instrumental to assure a balance between proper cytosine conversion and DNA degradation and undesired conversions that will avoid under- or over-estimation of methylation (Genereux *et al.*, 2008; *DNA Methylation Protocols*, 2018).

Methylation arrays: After bisulfite conversion, DNA is amplified and hybridized to arrays containing predesigned probes. Two types of probes are used to distinguish methylation status: one complementary to converted cytosines (unmethylated), and another complementary to the non-converted methylated cytosines (Yong *et al.*, 2016). Methylation is measured using Beta value, which estimates methylation level using a ratio of intensities between methylated and unmethylated alleles. Illumina® methylation arrays are the most widely used; by 2016 more than 360 publications used them (Kurdyukov and Bullock, 2016). Methylation arrays are focused on regions of interest based on the known functional effect of DNA methylation. Illumina`s latest implementation in the field is the MethylationEPIC Bead-Chip, which covers more than 850,000 CpG methylation sites, including 99% of RefSeq genes, 95% of CpG islands and high coverage of enhancer regions, ENCODE open chromatin, ENCODE transcription factor binding sites and miRNA promoter regions (Moran *et al.*, 2016). Thus, methylation arrays allow the study of methylation at single-base resolution in a cost-effective manner, with the main disadvantages being DNA degradation by bisulfite and the content-limited regions.

Bisulfite sequencing: Bisulfite sequencing is considered the gold-standard method in the study of DNA methylation. Workflow for WGBS (or MethylC-seq) library preparation includes DNA repair, adapter ligation, bisulfite treatment and polymerase chain reaction (PCR) amplification. One of the main disadvantages of bisulfite treatment is the DNA degradation it causes, but protocol improvements are dealing with this issue (*DNA Methylation Protocols*, 2018). In pre-bisulfite protocols, bisulfite treatment takes place after adapter ligation and

before PCR amplification, while in the post-bisulfite protocols bisulfite treatment is also used as fragmentation step and is followed by DNA repair, adapter ligation and PCR amplification.

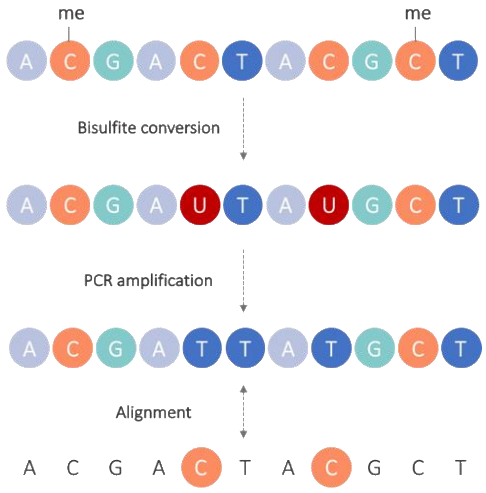

Figure 7. DNA bisulfite conversion. Bisulfite treatment causes deamination of unmethylated cytosines, converting them to uracils. In contrast, methylated cytosines remained unchanged. Once DNA has been treated with bisulfite, PCR amplification results in a non-complementary sequence, as methylated cytosines are not complementary to the PCR product that contains thymines at the corresponding positions. *A: Adenine, C: cytosine, G: Guanine, T: Thymine, U: Uracil, me: methylation mark*

The main advantages of post-bisulfite protocols are that degradation of the adaptor-tagged sequences caused by bisulfite treatment is avoided and the DNA input requirement is less than in the pre-bisulfite methods (*DNA Methylation Protocols*, 2018). Since the whole genome is targeted by this method, WGBS represents the most comprehensive approach for the assessment of methylation status in nearly every CpG site, including low-CpG-density regions. For the same reason, however, WGBS is the most costly method (Yong *et al.*, 2016). Because only a small fraction of the genome has the potential to be differentially methylated, **RRBS** has emerged as a tool for profiling DNA methylation at lower cost (Yong *et al.*, 2016). In this approach, DNA is fragmented by one or more restriction enzymes that cleave at sites containing a CpG locus. The two restriction enzymes most commonly used for RRBS are *Msp*I and *Taq*I, which cut into C↓CGG and T↓CGA, respectively. Digestion results in a mix of fragments of different sizes containing at least two CpG loci. Fragments are subjected to end repair, adapter ligation, size selection and bisulfite treatment before PCR amplification (*DNA Methylation Protocols*, 2018). Thus, in RRBS genome-wide coverage is obtained in areas of dense CpG at single-base resolution. However, this method may exhibit a lack of coverage at intergenic and distal regions and limits analysis to regions with the corresponding restriction motifs (Barros-Silva *et al.*, 2018).

One of the most challenging steps in bisulfite sequencing is mapping bisulfite-converted reads. After bisulfite treatment, reads are not complementary to the reference genome (Figure 7)

and every given thymine can be either an actual thymine or an unmethylated cytosine. Thus, special alignment tools are required for the alignment of bisulfite sequencing reads (Barros-Silva *et al.*, 2018; Yong *et al.*, 2016). Apart from single-base resolution, the main strength of bisulfite sequencing methods is the ability to quantify methylation levels. However, this quantification has been reported to be imperfect. In comparing different methods, Harris *et al*. reported a methylation level of ~18% of CpG varying by >25% between RRBS and WGBS, and the methylation level of ~5–8% of CpGs varying by >25% in RRBS biological replicates (Harris *et al.*, 2010).

2.3.1.2 Enriched-based methods

Although bisulfite sequencing is considered the gold standard method, methylation can also be profiled using NGS without bisulfite conversion. This section will be focus on enrichment methods for both unmethylated and methylated DNA.

Unmethylated DNA enrichment: Methylation-sensitive restriction enzyme sequencing (MRE-seq) consists of the sequencing of DNA fragmented using methylation-sensitive enzymes such as *Bst*UI, *Hpa*II, *Not*I and *Sma*I that cut only restriction sites when CpGs are unmethylated. Typically, DNA is digested with three to five methylation-sensitive restriction endonucleases and the resulting DNA fragments are size-selected and sequenced (*DNA Methylation Protocols*, 2018). MRE-seq provides DNA methylation estimation at single CpG resolution, but it is considered low-coverage due to the limit of enzyme recognition sites (*DNA Methylation Protocols*, 2018). Data analysis is performed by comparing the relative abundance of fragments (Yong *et al.*, 2016). Presence of the reads in a region is understood as unmethylated status of the corresponding CpG sites; however, lack of reads could not be inferred as a methylated site since perfect digestion is not typically achieved in practice (*DNA Methylation Protocols*, 2018). The combination of MRE-seq with methylated immunoprecipitation sequencing (MeDIP-seq) has been proposed as a method for comprehensive whole-methylome analysis (Harris *et al.*, 2010).

Methylated DNA enrichment: Methylated DNA can be enriched by immunoprecipitation, using specific antibodies against methyl-CpG (methylated immunoprecipitation, MeDiP), or by MBD proteins. The unmethylated fraction is typically discarded, and methylated fragments

can be profiled using microarray (MBD-chip, MeDIP-chip) (Nair *et al.*, 2011) or, more frequently, NGS (MBD-seq, MeDiP-seq). Apart from the protein used for pulling down methylated DNA, both protocols differ in input DNA and elution of methylated DNA. Immunoprecipitation of methylated DNA needs a denaturation step before incubation with the antibody, while MBD-enrichment is performed with double-stranded DNA. Once DNA is captured with the proteins it must be eluted: elution consists of a single step in MeDIP protocol, while MBD capture method allows single step or step-wise elution using increasing concentrations of salt. With the step-wise elution methylated DNA fragments can be stratified depending on their methylation density (Barros-Silva *et al.*, 2018; *DNA Methylation Protocols*, 2018).

Some differences in CpG densities of enriched DNA have been found between these methods. MeDIP enrichment preferentially captures regions of low CpG density (Harris *et al.*, 2010), which can be useful when methylation needs to be assessed in lower CpG regions, such as CpG shores or gene bodies. On the other hand, MBD enrichment offers the possibility of studying regions with different CpG densities: when single elution is performed a broad range of CpG densities is captured, while high CpG regions are obtained when high-salt buffers are used. Thus, MBD allows the study of methylation changes occurring in regions with low CpG density and regions with higher CpG density (Harris *et al.*, 2010), such as CpG islands. These differences are explained by the main differences in the protocols. In the MeDIP protocol DNA is denatured before immunoprecipitation. CpG dense fragments either re-anneal after denaturation or are not completely denatured and therefore do not bind effectively to the antibody. On the other hand, the affinity for methylated DNA varies depending on the protein used, so antibody enrichment may be more sensitive to fragments that contain single CpG sites (Nair *et al.*, 2011). Although enrichment enables improved coverage, it is sometimes considered as a weakness of the method because it also implies a bias toward hypermethylated areas (Yong *et al.*, 2016; Barros-Silva *et al.*, 2018). However, this issue can be resolved with a proper knowledge of the methods.

Methylation-enriched DNA sequencing is cost-effective and, since only a fraction of DNA is sequenced, read depth coverage is optimized and accuracy is improved (Barros-Silva *et al.*, 2018). Unlike microarrays and bisulfite sequencing, MBD-seq and MeDIP-seq do not provide single-base resolution. Extracted fragments can contain one or more methylated sites,

therefore the amount of methylation for a locus is estimated by the number of fragments covering that region (*DNA Methylation Protocols*, 2018). Moreover, as no change in sequence is introduced, alignment is easier than in bisulfite sequencing methods**.**

2.4 METHYLATION AND PREECLAMPSIA

DNA methylation changes have been described to occur in several diseases. With the aim of exploring if DNA methylation changes have been reported in preeclampsia a bibliography search was performed using PubMed; those studies carried out in humans that include preeclampsia (or pre-eclampsia) and methylation in their title or abstract were obtained. They were manually curated to selected those in which nuclear DNA methylation was analyzed in at least one tissue from women with preeclampsia, and significant differences were observed compared to control pregnancies. As a result, 49 studies were obtained. Among them, 33 are focused on the study of specific genes, while the remaining 16 followed a genome-wide approach.

Gene-specific methylation studies performed in preeclampsia so far (summarized in Table 2) are heterogeneous, not only regarding the gene of interest but also the subject of analysis. In the majority of them, placenta is the analyzed tissue, but some studies analyze maternal specimens (peripheral blood cells or omental arteries) or fetal ones (umbilical cord blood cells), focusing on methylation changes in women with preeclampsia and in their offspring, respectively. The diversity of the studies makes unified conclusion difficult, except that specific methylation changes have levels.

Table 2. Gene-specific methylation studies in preeclampsia

Authors, year	Tissue	Gene(s) of interest	Results
(Chelbi *et al.*, 2007)	Placenta	*SERPINA3*	CpG island methylation diminished in PE and growth restriction
(Wang *et al.*, 2010)	Placenta	*MMP9*	Less methylated in PE
(Gao *et al.*, 2011)	Placenta	*H19*	Hypermethylated in PE
(Mousa *et al.*, 2012c)	Omental arteries	*TBXAS1*	Hypomethylated in PE
(Mousa *et al.*, 2012b)	Omental arteries	*COL*, *MMP*, and *TIMP*	Hypomethylated in PE: *MMP1, MMP8, MMP12, MMP13, MMP13, MMP21, MMP26*. Hypomethylated in PE: *COL8A1, COL9A3, COL19A1, MMP16, MMP27, TIMP3*
(Founds *et al.*, 2012)	Placenta, maternal PBC	36 candidate genes	Hypermethylated in maternal cells and placenta: *ELL2, F11R*
(Xiang *et al.*, 2013)	Placenta, maternal PBC, UC-WBC	*TIMP3*	Hypomethylated in placenta PE
(He *et al.*, 2013)	UC-WBC	*IGF2, GNAS*	Hypomethylated in PE: *IGF2*
(Zhuang *et al.*, 2014)	Placenta	*Syncytin-1*	Hypermethylated in PE
(Doridot *et al.*, 2014)	Placenta	miR-34a	Hypomethylated in PE
(Ma *et al.*, 2014)	Placenta	*PEX1, GATAD1*	Hypomethylated in PE: GATAD1
(Lu *et al.*, 2014)	Placenta	*H19*	Hypermethylated in PE
(Liu *et al.*, 2014b)	Placenta	*MASPIN*	Hypomethylated in PE

Table 2. Gene-specific methylation studies in preeclampsia (Continuation)

Authors, year	Tissue	Gene(s) of interest	Results
(Hu *et al.*, 2014)	UC-WBC	*HSD11B2*	Hypomethylated in PE
(Than *et al.*, 2014)	Placenta, UC-WBC	*LGALS13, LGALS14, LGALS16*	Hypomethylated in PE: *LGALS16, LGALS14* Hypomethylated in PE: *LGALS13, LGALS14*
(Liu *et al.*, 2014a)	Placenta	*YWHAQ*	Hypermethylated in PE
(Anderson *et al.*, 2015)	Placenta	*CYP27B1, VDR, RXR*	Hypermethylated in PE
(Shimanuki *et al.*, 2015)	Placenta	*DLL1*	Hypermethylated in PE
(Tang *et al.*, 2015)	Placenta	*HLA-G*	Hypermethylated in PE
(Wilson *et al.*, 2015)	Placenta	*INHBA, PAPPA, FN1*	Hypermethylated in PE
(Ye *et al.*, 2016)	Placenta, maternal PBC	*GNA12*	Hypomethylated in PE
(White *et al.*, 2016)	Maternal PBC	33 candidate genes	Hypermethylated in PE: *CALCA, MTHFR, POMC, PTGS2*; Hypomethylated in PE: *AGT, DDAH1*
(Rahat *et al.*, 2016b)	Placenta, maternal PBC	Tumor suppressor genes	Hypermethylated in placenta PE: *PRKCDBP* Hypomethylated in placenta PE: *APC* Hypermethylated in maternal cells: *RASSF1A*
(Rahat *et al.*, 2017)	Placenta, maternal PBC	*VEGF, EGFR, c-JUN*	Hypermethylated in placenta PE: *VEGF*, *c-JUN* Hypermethylated in maternal blood cells: *c-JUN*
(Jia *et al.*, 2017)	Placenta	*IGFBP5*	Hypomethylated in PE
(Xiao *et al.*, 2017)	Placenta	*TFPI-2*	Hypomethylated in PE

Table 2. Gene-specific methylation studies in preeclampsia (Continuation)

Authors, year	Tissue	Gene(s) of interest	Results
(Sari *et al.*, 2017)	Maternal PBC	*EPHX2*	Hypomethylated in PE
(Liu *et al.*, 2017b)	Placenta	*WNT2*	Hypermethylated in PE
(Liu *et al.*, 2017a)	Placenta	*TF*	Hypomethylated in PE
(van den Berg *et al.*, 2017)	Placenta, UVEC, UC-WBC	Circadian clock and clock-controlled genes	EOPE vs spontaneous preterm: 6 CpGs in placenta, 21 CpGs in UC-WBC EOPE vs FGR: 2 CpGs in placenta and 8 CpGs in UC-WBC EOPE vs uncomplicated controls: 6 CpGs in placenta and 11 CpGs in UC-WBC
(Zhu *et al.*, 2018)	Placenta	*CD39, ZDHHC14*	Hypomethylated in PE
(Wang *et al.*, 2018)	UC-WBC	*MEST, DLK1*	Hypermethylated in PE: *MEST* (5 CpG sites) Hypomethylated in PE: *MEST* (2 CpG sites), *DLK1* (2 CpG sites) Mean methylation levels of the CpG sites were not significantly different
(Zhang *et al.*, 2019)	Placenta	*DKK1, WNT2*	Hypomethylated in PE: *DKK1*

PE: Preeclampsia, EOPE: Early-onset preeclampsia, LOPE: Late-onset preeclampsia, FGR: Fetal growth restriction, PBC: Peripheral blood cells, UB-WBC: Umbilical blood white blood cells, UVEC: umbilical vein endothelial cells

Among genome-wide methylation studies in preeclampsia (Table 3), most assess placenta (n=11). Methods are also homogeneous; 12 out 16 studies use methylation array alone, and 2 out of 16 use them in combination with immunoprecipitation (MeDIP-chip). However, the lack of standards for analyzing large datasets for methylation arrays means the criteria to filter differentially methylated probes (DMPs) vary considerably between studies. Furthermore, the algorithm used for the analysis of array data differs between studies. Chu *et al.* demonstrated that, using the same data, different statistically significant CpG sites are identified when different algorithms are used, highlighting the influence of statistically significant regions (Chu *et al.*, 2014). Zhao *et al.* (Zhao *et al.*, 2017) performed their study using raw data from another study (Martin *et al.*, 2015). The use of different software and cut-offs resulted in differences in the DMRs found (Table 3). The absence of correlation can be also explained by preeclampsia itself, because clinical characteristics can vary widely between women as can the preeclampsia classifications used in the studies (early-onset preeclampsia, late-onset preeclampsia, preterm preeclampsia, term preeclampsia).

Despite the lack of correlation, the main conclusion that can be drawn from all of these studies is that methylation changes are associated with preeclampsia, not only at placental level, but also at maternal and fetal levels.

Then, the discovery of disease-specific biomarkers based on DNA methylation changes represents an interesting approach in preeclampsia study. However, the search for methylation differences may be best realized through minimally invasive approaches that do not require analysis of placental tissue.

Table 3. Genome-wide methylation studies in preeclampsia

Authors, year	Groups of study (n)	Tissue	Method	Number of significant DMRs
(Yuen *et al.*, 2010)	EOPE (4), LOPE (4), IUGR (4), Early control (4), Late control (5), Control (5)	Placenta	Illumina GoldenGate Methylation Cancer Panel I	192 EOPE, 16 in IUGR, 0 in LOPE
(Jia *et al.*, 2012)	PE (9), Control (9)	Placenta	MeDIP-chip (NimbleGen)	296 genes: 102 hypermethylated, 194 hypomethylated
(Wright *et al.*, 2012)	PE (6), Control (6)	Placenta, maternal PBC	Array (unknown)	Maternal blood cells: 123 hypermethylated and 67 hypomethylated Placenta: 63 hypermethylated and 69 hypomethylated
(Mousa *et al.*, 2012a)	PE (7), Control (7)	Maternal omental arteries	Infinium array 27K	FDR < 10%: 1685 genes, FDR < 5%: 236 genes
(Blair *et al.*, 2013)	EOPE (20), Control (20)	Placenta	Illumina array 450K	282 DMPs: 72 hypermethylated, 210 hypomethylated
(White *et al.*, 2013)	PE (14), Control (14)	Maternal PBC (delivery)	Infinium array 27K	997 DMPs: 729 hypermethylated, 268 hypomethylated
(Anderson *et al.*, 2014)	PE (6), Control (6)	Placenta, Maternal PBC (1st trimester)	Infinium array 27K	Maternal PBC: 132 hypermethylated and 75 hypomethylated. Placenta Most of the regions found in maternal PBC, with varying degrees of methylation

Table 3. Genome-wide methylation studies in preeclampsia (Continuation)

Authors, year	Groups of study (n)	Tissue	Method	Number of significant DMRs
(Anton *et al.*, 2014)	PE preterm (12), PE term (19), Control term (14)	Placenta	Illumina array 450K	Term + Preterm PE vs Control: 0 DMPs Preterm PE vs Control, 421 DMPs: 396 hypermethylated, 25 hypomethylated Term PE vs Control, 0 DMPs
(Chu *et al.*, 2014)	PE (24), Control (24)	Placenta	Infinium array 27K	DNA methylation is strongly associated with fetal gender. 49 DMPs: 11 hypermethylated, 38 hypomethylated
(Zhu *et al.*, 2015)	LOPE (4), Control (4)	Placenta	MeDIP-seq and hMeDIP-seq	714 DMRs: 487 hypermethylated, 227 hypomethylated
(Martin *et al.*, 2015)	PE (19), Control (17)	Placenta	HumanMethylation450k BeadChip array	989 DMPs: 191 hypermethylated, 798 hypomethylated
(Xuan *et al.*, 2016)	PE (12), Control (12)	Placenta	MeDIP-chip (NimbleGen)	1664 promoters: 663 hypermethylated, 1001 hypomethylated
(Yeung *et al.*, 2016)	PE (8), Control (16)	Placenta	Illumina HumanMethylation450K BeadChip	303 DMRs: 214 hypermethylated, 89 hypomethylated

Table 3. Genome-wide methylation studies in preeclampsia (Continuation)

Authors, year	Groups of study (n)	Tissue	Method	Number of significant DMRs
(*Suzuki et al., 2016*)	PE (36), Hypertension (11), Control (15)	Placenta	HELP-tagging	__
(Herzog *et al.*, 2017)	EOPE (13), LOPE (16), Control (36), Normotensive FGR (27), Normotensive preterm (20)	Placenta, UB-WBC, UVEC	Illumina HumanMethylation450K BeadChip	EOPE vs preterm Control: 5001 in UC-WBC, 869 in placenta
(Zhao *et al.*, 2017)	PE (19), Control (17)	Placenta	Illumina HumanMethylation450K BeadChip	2667 DMRs: 1433 hypermethylated, 1234 hypomethylated

PE: Preeclampsia, EOPE: Early-onset preeclampsia, LOPE: Late-onset preeclampsia, IUGR: Intrauterine growth restriction, FGR: Fetal growth restriction, PBC: Peripheral blood cells, FDR: False discovery rate, UB-WBC: Umbilical blood white blood cells, UVEC: umbilical vein endothelial cells, DMPs: Differentially methylated probes, DMRs: Differentially methylated regions

3. CIRCULATING CELL-FREE DNA

Circulating cell-free DNA (cfDNA) refers to fragmented DNA present in the non-cellular component of the blood (Corcoran and Chabner, 2018). In 1948, 5 years before the discovery of the structure of DNA, Mandel and Metais discovered the presence of circulating cell-free DNA and RNA in human blood of healthy and sick individuals (Mandel and Metais, 1948). In that moment, no one could imagine that this discovery would represent the first step to a minimally invasive revolutionary tool: liquid biopsy. However, this tool has been a long time in coming. During the decade following discovery of cfDNA, no interest was shown for circulating nucleic acids research. It was not until the mid-1960s when Tan *et al.* discovered high levels of cfDNA in patients with systemic lupus erythematosus (Tan *et al.*, 1966). Again, progress languished until increased concentrations of cfDNA were detected in patients with cancer compared to healthy individuals (Leon *et al.*, 1977), and another twelve years later, several cfDNA molecules in plasma of cancer patient were identified to have neoplastic characteristics (Stroun *et al.*, 1989). In 1994, two works supported this finding and established that cfDNA in cancer patients carried tumor-associated oncogene mutations, evidencing that tumor cells release their DNA into the circulation (Vasioukhin *et al.*, 1994; Sorenson *et al.*, 1994). These works paved the way to clinical translation, showing the promising use of cfDNA in tumor detection and monitoring, though sufficiently sensitive and specific techniques had yet to be developed (Volik *et al.*, 2016). The discovery of circulating tumor DNA coexisting with healthy cells motivated further study of cfDNA. In 1997, Lo *et al.* discovered the presence of fetal cfDNA in the plasma of pregnant women (Lo *et al.*, 1997), opening up the field to the development of fetal cfDNA-based aneuploidy testing (non-invasive prenatal testing, NIPT).

3.1 CHARACTERISTICS OF CELL-FREE DNA

Circulating cfDNA is a mix of DNA molecules originating from different parts of the body, with hematopoietic cells the main contributor (Moss *et al.*, 2018; Sun *et al.*, 2015). In healthy individuals cfDNA concentration ranges between 1 and 10 ng/mL (Wan *et al.*, 2017), with increased levels associated with cancer and pregnancy, but also with physical exercise, inflammation, diabetes, tissue trauma or sepsis (Kustanovich *et al.*, 2019). In pregnant women, fetal cfDNA found in maternal plasma originates from placental cells (Bianchi *et al.*, 2015). cfDNA is mainly composed of double-stranded fragments of nuclear DNA and, to a lesser

extent, mitochondrial DNA (mtDNA) (Kustanovich *et al.*, 2019). Fragment size depends on their origin; circulating mtDNA is short in plasma [< 100 base pairs (bp)] (Zhang *et al.*, 2016; Jiang *et al.*, 2015). Nuclear cfDNA presents a characteristic peak at 166 bp, which corresponds to the size of DNA wrapped around one nucleosome (~147 bp) plus linker DNA associated to histone H1 (Wan *et al.*, 2017). This peak is followed by two smaller peaks around 300 and 450 bp, corresponding with linear progression of nucleosome units (Volik *et al.*, 2016) (Figure 8). cfDNA fragment size profiles show a 10-bp ladder pattern, which is presumably caused by nuclease cleavage of exposed sites corresponding to each turn of the DNA helix (Wan *et al.*, 2017; Lo *et al.*, 2010). Apart from the differences due to genomic origin (nuclear or mitochondrial), differences in cfDNA size are also observed by tissue of origin. Non-hematopoietically derived DNA is shorter than DNA from hematopoietic cells, presenting a reduction in the 166-bp peak (Zheng *et al.*, 2012). As plasma DNA is predominantly hematopoietic in origin, most of the fragments contribute to the mononucleosomal peak. An increase in non-hematopoietic cfDNA results in an increase in shorter fragments.

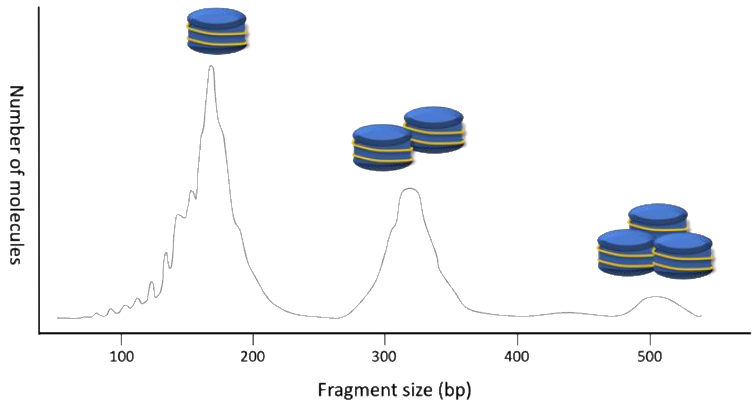

Figure 8. Schematic representation of cfDNA fragment distribution. Size of DNA molecules found in human circulation is characterized by linear progression of nucleosome units. Most of the molecules are around 166 bp, the size of DNA wrapped around one nucleosome (~147 bp) plus linker DNA associated to histone H1. This peak is followed by two smaller peaks around 300 bp and 450 bp, which correspond to 2- and 3-nucleosome sizes. Peaks before the mono- and di-nucleosome peaks follow a 10-bp periodicity and are presumably caused by nuclease cleavage of sites exposed corresponding to each turn of the DNA helix.

There are several sources of cfDNA: due to cell death either by necrosis, apoptosis or pyroptosis, and due to active mechanisms of cfDNA release from live cells (Aucamp *et al.*, 2018). Of them, many studies suggest that the main source of cfDNA in both healthy and sick

individuals is apoptosis (Aucamp *et al.*, 2018). The fact that cfDNA fragment size correlates with the sizes of caspase-dependent cleavage supports this idea (Kustanovich *et al.*, 2019).

The half-life of cfDNA molecules in blood circulation is short, from several minutes to 1-2 hours (Kustanovich *et al.*, 2019). Pregnancy represents an ideal scenario to study cfDNA clearance, since the moment in which the source of fetal cfDNA disappears is certain. Fetal cfDNA disappears in two stages: the half-life of cfDNA cleared in the first rapid phase is 1-2 h, while that cleared in the second slow phase is around 13 h. After 24-48 h post-delivery, no fetal cfDNA remains in maternal blood (Yu *et al.*, 2013).

3.2 CELL-FREE DNA AND EPIGENETICS

Plasma cfDNA represents a wide source of information. The most obvious one is DNA sequence, which allows the identification of mutation from cancer cells (Stroun *et al.*, 1989), fetal DNA using single-nucleotide polymorphisms (Lo *et al.*, 2010) or donor-specific DNA in transplantation recipients (Lo *et al.*, 1998). Nevertheless, it is not the only information cfDNA carries: epigenetics in cfDNA is emerging as a promising source of biomarkers in conditions that do not include DNA mutations.

3.2.1 DNA methylation in cell-free DNA

The existence of tissue-specific methylation signatures (Fernandez *et al.*, 2012; Kundaje *et al.*, 2015) has been used to trace the tissue of origin of cfDNA present in plasma (Sun *et al.*, 2015; Lehmann-Werman *et al.*, 2016; Moss *et al.*, 2018). Studying the tissue of origin in cancer is interesting because an increase in the cfDNA corresponding to the tissue in which the tumor is present is observed in cancer patients (Sun *et al.*, 2015; Lehmann-Werman *et al.*, 2016; Moss *et al.*, 2018). Its potential in the early diagnosis of cancer has been demonstrated in cases of incidental detection of occult maternal malignancies during NIPT (Bianchi *et al.*, 2015). In non-cancerous diseases, tissue of origin tracing has been successfully used in different diseases and conditions such as type 1 diabetes and pancreatic islet transplantation (Lehmann-Werman *et al.*, 2016; Moss *et al.*, 2018), bone marrow transplantation (Sun *et al.*, 2015), sepsis (Moss *et al.*, 2018), multiple sclerosis, traumatic brain injury, cardiac arrest and pancreatitis (Lehmann-Werman *et al.*, 2016). In the case of pregnancy, cfDNA analysis allows the study of the placental contributions to maternal plasma and has been used to reconstruct the placental

methylome from maternal plasma cfDNA (Sun *et al.*, 2018b). This presents a promising, minimally invasive opportunity to assess placental health.

The most widespread use of cfDNA methylation study has been the identification of changes between healthy and sick individuals. Methylation changes associated with disease have been proven most obvious in cancer (Fernandez *et al.*, 2012), and research on discovery of methylation biomarkers by mining the methylome of cancer cells is the most abundant (Feng *et al.*, 2018; Gai and Sun, 2019).

Disease is associated with changes in the proportions of cfDNA originating from different tissues and also with DNA methylation changes, so cfDNA methylation analysis provides two types of information: tissue of origin of cfDNA and disease-specific methylation changes. When DMRs are found between healthy and sick individuals, they can be due to changes in tissue contribution or to DNA methylation changes in particular tissues (Feng *et al.*, 2018).

3.2.2 Size of cell-free DNA and nucleosome positioning

Circulating cfDNA molecules are not randomly fragmented: DNA fragments that are released into circulation are those that are protected from nuclease degradation by bound proteins. These proteins are typically histones and transcription factors, which explains the nucleosome patterns size observed in cfDNA (Volik *et al.*, 2016). Based on that, research is starting to focus on the information that cfDNA fragmentation can provide.

On one hand, cfDNA molecule size analysis has emerged as a new field of study in cancer and NIPT, since it was demonstrated that there are size differences in molecules originating from different tissues (Zheng *et al.*, 2012). Indeed, placental cfDNA found in maternal plasma during pregnancy shows a reduction of the 166-bp fragment and a relative prominence around 143 bp (Lo *et al.*, 2010). The reason for this difference is that nucleosome accessibility in placental cells is higher than in white blood cells (Sun *et al.*, 2018a). The same occurs in cancer cfDNA, where tumor cfDNA is more abundant in the shorter fraction of circulating DNA (Mouliere *et al.*, 2018; Jiang *et al.*, 2015; Hellwig *et al.*, 2018). This difference is being used to enhance the detection of tumor cfDNA (Hellwig *et al.*, 2018; Mouliere *et al.*, 2018) as well as to enrich fetal cfDNA in NIPT (Yu *et al.*, 2014).

On the other hand, nucleosome occupancy is highly related to cell identity. Based on that, Snyder *et al.* developed a strategy to identify the tissue of origin of cfDNA. The authors hypothesized that tissue of origin could be inferred because cfDNA molecules might contain evidence of the epigenetic landscape from which the cfDNA derives. With this method, they presented a genome-wide map of in vivo nucleosome protection inferred from cfDNA fragments and were able to trace the tissue of origin of cfDNA (Snyder *et al.*, 2016). Along the same lines, Ulz *et al.* analyzed cfDNA nucleosome occupancy at promoter regions. The existence of an association between actively-transcribed genes and nucleosome-depleted promoters motivated them to investigate if cfDNA is able to reflect such expression-specific nucleosome occupancy at promoters. Whole-genome sequencing of cfDNA showed that differences in nucleosome occupancy between expressed and silent genes results in different read depth coverage patterns (Ulz *et al.*, 2016) and demonstrated its potential use to infer the tissue of origin of cfDNA.

Fragmentation of cfDNA is not random, but not only because of the presence of proteins that protect DNA from degradation. Recent studies showed that the cutting sites found in cfDNA molecules are not randomly chosen. Analysis in pregnant women showed that certain genomic locations were overrepresented at the ends of plasma DNA fragments with fetal or maternal selectivity and demonstrated that cfDNA molecules exhibit preferred end sites depending on whether their origin is fetal or maternal (Chan *et al.*, 2016). Taking together cut and size differences between maternal and fetal cfDNA, further analysis was carried out by Sun et al. Circulating DNA molecules were divided into short and long depending on their size, and preferred ends of molecules were analyzed. Maternal DNA was mostly cut within the linker region, while fetal DNA was frequently cut within the nucleosome core, demonstrating higher nucleosome accessibility in the placental cells. Similar differences were also observed in non-pregnant healthy subjects, suggesting that the interrelationship of plasma DNA size and preferred DNA ends is likely a general one, extending its analysis outside the context of pregnancy (Sun *et al.*, 2018a).

3.3 CELL-FREE DNA AND PREECLAMPSIA

In 1999, Lo *et al.* described for the first time an increase in fetal cfDNA in women with preeclampsia compared to control pregnancies (Lo *et al.*, 1999). Since that moment, study of

cfDNA in preeclampsia has been limited to studying its concentration, and focus has been put on the use of fetal cfDNA quantification as a predictor of the future development of preeclampsia. Higher levels of fetal cfDNA in preeclampsia have been attributed mainly to an increase in apoptosis of trophoblastic cells due to the hypoxic conditions within the intervillous space that cause placental ischemia (Martin *et al.*, 2014). Impaired fetal DNA clearance from maternal plasma could also contribute to that increase (Lau *et al.*, 2002).

Martin *et al.* performed a systematic review in which they found that 11 out of 13 studies measuring fetal cfDNA by qPCR showed a significant increase in fetal cfDNA in preeclampsia that was more evident in the cases of severe or early-onset preeclampsia than in mild or late-onset cases (Martin *et al.*, 2014). A meta-analysis by Contro *et al.* analyzed published data and suggested that quantification of fetal cfDNA by qPCR for the prediction of preeclampsia could be performed from the beginning of the 2nd trimester, and not before (Contro *et al.*, 2017). However, in both reviews the heterogeneity between the reviewed studies did not enable conclusions about the statistical and clinical relevance (Martin *et al.*, 2014; Contro *et al.*, 2017). Definition of preeclampsia, timing of assessment of cfDNA and markers used for fetal cfDNA were not consistent between studies. Moreover, sample size of the studies was limited, and disease classification varied: some used early-onset or severe preeclampsia, while others used any type of preeclampsia (Martin *et al.*, 2014).

More recent studies have analyzed cfDNA in preeclampsia in larger groups using NGS as a part of prenatal screening. In these studies, fetal cfDNA is not measured as an absolute quantification, instead it is measured as the amount of fetal cfDNA detected in the plasma compared to the total cfDNA, called fetal fraction. Rolnik *et al.* analyzed plasma samples at 11-13 weeks of gestation and found an increase in total cfDNA but a reduction in fetal fraction in women who subsequently developed early preeclampsia. However, after the adjustment for maternal characteristics that affect both total cfDNA and fetal fraction measurements, differences were not significantly different from those in normotensive controls (Rolnik *et al.*, 2015). The same was observed in a study that analyzed 1st-trimester samples in women who later developed gestational hypertension, preeclampsia and severe preeclampsia. Although the fetal fraction was lower in the women who developed any hypertensive disorders compared to controls, and severe preeclampsia had the lowest value, the differences were

not statistically significant after adjusting for potential confounders (Bender *et al.*, 2019). A different approach was used by Rolnik *et al.*; correlation between fetal fraction and biophysical markers for preeclampsia was analyzed in 4317 women. Interestingly, they found that the lower the fetal fraction, the higher the risk for preeclampsia. However, outcomes of the pregnancies were not collected, thus the association was made only with the estimated risk and not with real incidence. The authors remark that the capacity of fetal fraction measurement to act an as independent 1st-trimester marker in an algorithm screening requires further research (Rolnik *et al.*, 2018). Along the same lines, a study that examined fetal fraction in cases of placental compromise and adverse perinatal outcomes found that low fetal fraction was associated with placental compromise, pregnancy hypertension and severe preeclampsia (Gerson *et al.*, 2019).

These works have provided evidence of the potential use of cfDNA analysis in preeclampsia. Widening cfDNA analysis by exploring some of the several types of information cfDNA can provide, such as epigenetics, would represent a potential tool for disease-associated changes identification.

II. HYPOTHESIS

II. HYPOTHESIS

The hypothesis of this thesis is that cell-free DNA from pregnant women with severe preeclampsia exhibits differential methylation compared to that from women with control pregnancies. Detecting such differences will facilitate further studies to delineate underlying pathophysiology and discover novel disease biomarkers for preeclampsia diagnosis and prognosis.

III. OBJECTIVES

III. OBJECTIVES

- To establish a method to detect changes in cell-free DNA methylation status in a clinical diagnostic laboratory workflow.

- To compare cell-free DNA methylomes in severe preeclampsia versus control pregnancies searching for differentially methylated regions.

- To uncover the role of methylation changes in severe preeclampsia physiopathology and symptoms.

- To investigate the utility of other cell-free DNA epigenetic features, such as fragmentation patterns, for biomarker identification in severe preeclampsia.

IV. MATERIAL & METHODS

IV. MATERIAL AND METHODS

1. OVERVIEW

To study differences in methylation in cfDNA between severe preeclampsia (sPE) and control pregnancies, blood samples were collected from 41 pregnant women with and without preeclampsia diagnosis. Plasma was obtained from blood centrifugation, and 1 mL from each sample was used to extract cfDNA. After cfDNA incubation with paramagnetic beads coupled to methyl-CpG binding domain of human MBD2 protein, stepwise elution was performed to obtain methylated cfDNA fraction. Methylated cfDNA was purified and used to prepare a DNA library (Figure 9)

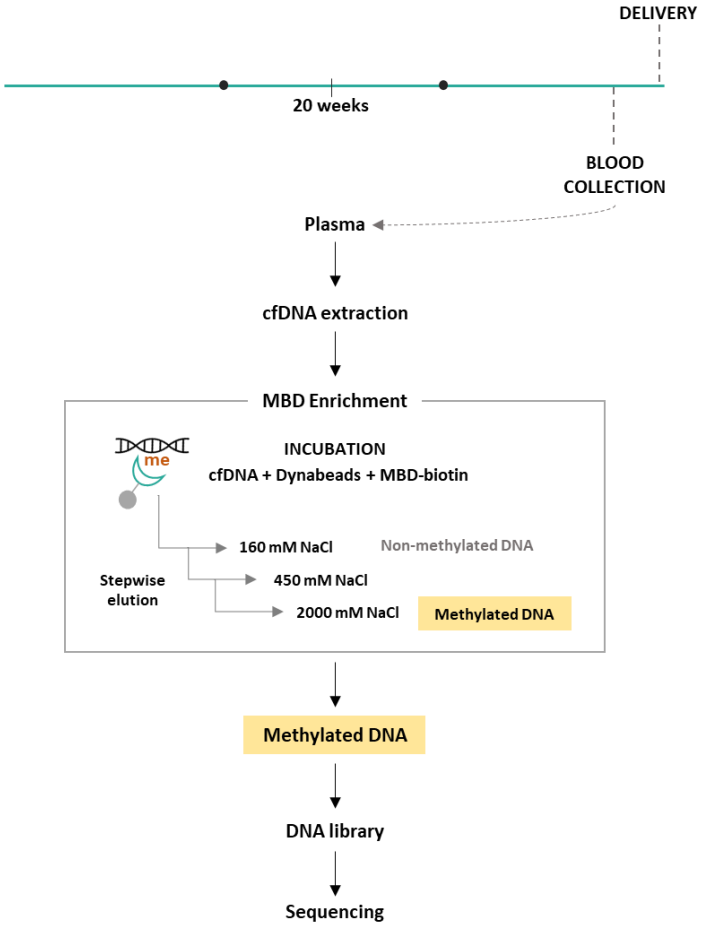

Figure 9. Study design summary. Cell-free DNA was extracted from blood of pregnant women with and without severe preeclampsia diagnosis. Methylated fraction was obtained by MBD-enrichment and sequenced using next generation sequencing.

Two sequencing runs were performed using an Illumina NextSeq 500 sequencer. The first one, an exploratory run, was performed using single-read sequencing with 24 samples (12 per

group). Then, sample size was increased by 17 additional specimens (Control n=22, sPE n=19). Samples were pooled and paired-end sequencing was carried out with the aim of obtaining a deeper knowledge as well as to perform fragment size analysis.

2. INCLUSION/EXCLUSION CRITERIA

This study was approved by the Clinical Ethics Committee at Hospital Univeristario y Politécnico La Fe (Spain) (2011/0383), and all participants provided informed consent.

Women attending Hospital La Fe were classified as control pregnancies when they were not diagnosed with any obstetric complication. Women were diagnosed as severe preeclampsia when they presented elevated blood pressure (systolic pressure ≥160 mm Hg or diastolic pressure ≥ 90 mm Hg) or any of the following: proteinuria, thrombocytopenia, impaired liver function, progressive renal insufficiency, edema, or the new onset of cerebral or visual disturbances. Women meeting these criteria were excluded if they arrived at the hospital in active delivery or presented with any condition that could cause the passage of fetal material into the maternal blood (hemorrhage, amniocentesis or chorionic villus sampling). Women were also excluded for other gestational complication such as placenta accreta, percreta or increta; gestational diabetes; fetal death or spontaneous abortion; malformed fetus; or cervical cerclage. Inclusion criteria comprised maternal age between 18 and 42 years, singleton pregnancy, gestational age between 23 and 42 weeks and complete follow-up until delivery. The clinical features of the participants are summarized in Tables 4 and 5.

Table 4: Demographic and clinical characteristics for women with samples in Run 1

	Control (N=12)	sPE (N=12)	p-value*
Maternal age (years)	31.9 (1.86)	33.6 (1.77)	ns
Systolic blood pressure (mmHg)	116.9 (4.03)	168.1 (3.73)	< 0.001
Diastolic blood pressure (mmHg)	69.9 (2.47)	102.8(2.87)	< 0.001
Proteinuria (mg)	0 or NA	+ 1 to +2	-
sFLT1/PIGF Ratio	0 or NA	201 (39.05)	-
Gestational age at sample collection (weeks)	32.8 (0.76)	32 (1.16)	ns
Gestational age at delivery (weeks)	33.1 (0.73)	32.4 (1.17)	ns
Sample collection (days before delivery)	2.6 (0.57)	3.1 (0.43)	ns
Birth weight (g)	2706.9 (186.6)	1770.1 (259.9)	< 0.05
Fetal sex (% male)	5 (41.7 %)	5 (41.7 %)	ns
Parity (n)	1.75 (0.18)	1.75 (0.22)	ns

Mean (Standard Error of the Mean)

NA: Not available

Proteinuria +1 = 30 to 100 mg/dL, +2 = 100 to 300 mg/dL

* Fetal sex = Chi-squared test; Other categories = One-tailed Student´s t-test

Table 5: Demographic and clinical characteristics for women with samples in Run 2

	Control (N=22)	sPE (N=19)	p-value*
Maternal age (years)	31.2 (1.45)	32.6 (1.34)	ns
Systolic blood pressure (mmHg)	119.1 (3.4)	165.1 (2.98)	< 0.001
Diastolic blood pressure (mmHg)	71.6(2)	101.9 (2.11)	< 0.001
Proteinuria (mg)	0 or NA	+1 to +2	-
sFLT1/PIGF Ratio	0 or NA	217.2 (60.06)	-
Gestational age at sample collection (weeks)	34.3 (1.02)	32.3 (0.86)	ns
Gestational age at delivery (weeks)	34.5 (1.01)	32.8 (0.89)	ns
Sample collection (days before delivery)	2.16 (0.44)	3.1 (0.58)	ns
Birth weight (g)	2436 (195.79)	1774.5 (183.1)	< 0.05
Fetal sex (% male)	11 (50 %)	8 (42.1 %)	ns
Parity (n)	2.04 (0.26)	1.84 (0.18)	ns

Mean (Standard Error of the Mean)

NA: Not available

Proteinuria +1 = 30 to 100 mg/dL, +2 = 100 to 300 mg/dL

* Fetal sex = Chi-squared test; Other categories = One-tailed Student´s t-test

3. BLOOD PROCESSING AND CELL-FREE DNA EXTRACTION

Peripheral blood samples (10 mL) were collected into cell-free DNA BCT collection tubes (Streck, NE, USA) and were centrifuged 15 min, 1600 rcf at 4ºC. Supernatant was centrifuged 10 min, 16000 rcf at 4ºC to separate plasma. Circulating cfDNA was extracted from 1 mL of plasma in QIAsymphony SP equipment using a modified protocol of QIAsymphony DSP Virus/Pathogen Kit (QIAGEN, Hilden, Germany). RNA Carrier was used to improve extraction yield, and cfDNA was eluted in a final volume of 50 µL.

4. METHYL-BINDING DOMAIN ENRICHMENT

MethylMiner Methylated DNA Enrichment Kit (Invitrogen, Carlsbad, USA) was used to isolate methylated DNA from 50 µL of extracted cfDNA. Buffer 1X Bind/Wash was prepared by diluting 5X Bind/Wash Buffer with DNAse-free water.

For 24 samples, 3 tubes containing 20 µL of Dynabeads M-280 Streptavidin (2.5 µL per sample) were washed twice; 180 µL of 1X Bind/Wash was added and mixed by pipetting. Tubes were placed on the magnetic rack; after 1 minute, liquid was removed and discarded with a pipette without touching the beads. Tubes were removed from the magnetic rack and 200 µL of 1X Bind/Wash was added to Dynabeads and mixed by pipetting. An additional wash was performed following the same steps.

Washed Dynabeads were coupled with methyl-CpG binding domain of human MBD2 protein MBD-Biotin Protein via a biotin linker. Three tubes containing 12.5 µL (1.5 µL per sample) MBD-biotin protein were diluted with 187.5 µL of 1X Bind/Wash. Each sample was mixed with one of the tubes of washed beads on a rotating mixer at room temperature for 1 hour. Tubes were then placed on the magnetic rack for 1 minute and liquid was removed and discarded by pipette without touching the beads. Tubes were removed from the magnet, 320 µL of 1X Bind/Wash were added and beads were mixed at room temperature for 5 minutes on a rotating mixer. Two additional washes were performed following the same steps. Contents were divided into 8 tubes: 40 µL of MBD-Beads were added to 50 µL of cfDNA and 10 µL of 5X Bind/Wash Buffer. cfDNA was incubated with MBD-beads on a rotating mixer overnight at 4ºC. After incubating the DNA and MBD-beads, stepwise elution was performed with 160 mM NaCl, 450 mM NaCl and 2000 mM NaCl buffers (Figure 9). Tubes were placed on the magnetic

rack and supernatant was removed by pipette without touching the beads. The supernatant, which contained the non-captured cfDNA, was saved and stored on ice. Beads were washed of residual non-captured DNA as follows: 100 µL of 1X Bind/Wash (160 mM NaCl) were added and tubes were placed on a rotating mixer for 3 minutes. With the tubes on the magnet, liquid was removed and store on ice. An additional wash step was performed following the same steps. Elution buffer (450 mM NaCl) was prepared by combining low-salt buffer (0 mM NaCl) and high-salt buffer (2000 mM NaCl) following kit instructions. After removing the non-captured DNA, two elution steps were performed using 450 mM NaCl buffer to remove cfDNA molecules with a small amount of methylated CpG sites and the possible non-methylated DNA. Briefly, 100 µL of 450 mM NaCl buffer were added to each tube; after 3 minutes on a rotating mixer, tubes were placed on the magnet and supernatant was removed and stored on ice. Immediately, two elution steps with high-salt buffer were performed to release methylated DNA molecules from the beads: 100 µL of 2000 mM NaCl buffer were added to each tube and mixed on a rotating mixer for 3 minutes. Once tubes were placed on the magnetic rack, supernatant was removed and stored on ice. After the two elution steps with high-salt buffer, the eluted 200 µL of methylated cfDNA were cleaned using Agencourt AMPure XP (Beckman Coulter, Brea, USA). 200 µL of AMPure beads were mixed and incubated for 5 minutes at room temperature with 200 µL of methylated cfDNA. Tubes were placed on the magnetic rack. After 5 minutes, liquid was removed and discarded. Immediately, two wash steps were performed: 200 µL of fresh ethanol 70% were added to each tube and removed after 30 seconds. After removing ethanol for the second wash, tubes remained on the magnet and were incubated at room temperature with the lid open for 5 minutes to air-dry. Tubes were removed from the magnetic rack and beads were resuspended in 60 µL of Buffer EB (QIAGEN, Hilden, Germany). Tubes were incubated at room temperature for 5 minutes and placed on the magnet for 2 minutes. Then, 50 µL of eluted DNA were removed without touching the beads.

5. DNA LIBRARY PREPARATION, POOLING AND LOADING

Sample indexing and library preparation for sequencing were performed from cleaned-up methylated fractions using TruSeq Nano DNA Library Prep Kit (Illumina Inc., San Diego, USA).

From each sample, 48 µL of methylated cfDNA were placed in wells of a 96-well plate. Then, 12 µL of End Repair Mix were added with mixing by vortex, and the plate was centrifuged at 280 x g for 1 minute. Samples were incubated in the thermal cycler 10 minutes at 30ºC, followed by 5 minutes at 75ºC and a final hold at 10ºC. A-Tailing mix and Ligation mix were diluted following kit instructions. Then, 15 µL A-Tailing mix were added to each sample well, followed by vortex and centrifugation. Plate was incubated 10 minutes at 37ºC, 5 minutes at 75ºC and a final hold of 10ºC. On a different 96-well plate, for each sample 18.6 µL of Ligation Mix were mixed with 1.5 µL of a different DNA Index Adapter. Once A-tailing was finished, 8 µL of the Ligation Mix + Adapter were added to each sample. The plate was incubated in the thermal cycler 10 minutes at 30ºC, 15 minutes at 75ºC followed by a final hold at 10ºC. Ligation products were purified using 83 µL of Sample Purification Beads. After incubation at room temperature for 5 minutes, the plate was placed on the magnetic stand for 5 minutes; then, supernatant was removed and discarded. Immediately, two washing steps with 200 µL of fresh 70% ethanol were performed. The plate was maintained on the magnet at room temperature for 7 minutes to let beads air-dry. Then, 30 µL of EB Buffer (QIAGEN) were added to beads, which were resuspended and incubated at room temperature for 2 minutes. The plate was placed on the magnetic stand for 5 minutes. Then, 25 µL of the supernatant were transfer to another plate.

PCR mix was prepared following kit instructions (400 µL of Enhanced PCR Mix were mixed with 100 µL of Primer PCR Cocktail Mix) and 25 µL of this mix was added to each sample well. PCR amplification consisted of 3 minutes at 95ºC; 13 cycles of 98ºC for 20 seconds, 60ºC for 15 seconds and 72ºC for 30 seconds; 5 minutes at 72ºC and final hold at 10ºC. After PCR enrichment, clean-up was performed using 50 µL of Sample Purification Beads added to each sample. After room-temperature incubation for 5 minutes, the plate was placed on the magnet for 5 minutes, then supernatant was discarded. After two washes with 200 µL fresh 70% ethanol, the plate was maintained on the magnet at room temperature for 7 minutes to let beads air-dry. Then, 120 µL of EB Buffer (QIAGEN) were added to beads, which were resuspended and incubated at room temperature for 2 minutes. The plate was placed on the magnetic stand for 2 minutes, then 100 µL of DNA library were removed and placed on another plate.

Libraries were quantified using Qubit dsDNA HS Assay kit (Invitrogen) following kit instructions and were diluted to 10 nM. Then, all samples were pooled and Buffer EB (QIAGEN) was added to reach a final concentration of 1.7 nM. Libraries were denatured and loaded into the sequencer (NextSeq 500 sequencing system, Illumina Inc., San Diego, USA). Single-end sequencing was performed for 150 cycles for the first group of samples and paired-end sequencing was performed for 2 x 100 cycles for the second group of samples.

6. DATA ANALYSIS

6.1 ALIGNMENT AND QUALITY FILTERING

Raw data from Illumina sequencing was downloaded from Illumina BaseSpace. A first quality control was performed using FASTQ files obtained using bcl2fastq software from Illumina. Then, each sample was aligned to the human reference genome (hg19) using BWA (http://bio-bwa.sourceforge.net/) (Li and Durbin, 2009). Reads with mapping quality lower than 10 were filtered out using Samtools (http://samtools.sourceforge.net/) (Li, 2011) and duplicates were removed with PICARD software (https://broadinstitute.github.io/picard/index.html).

6.2 CHROMOSOME Y CELL-FREE DNA QUANTIFICATION

A bigwig file was obtained from each sample after bam files were divided into 100 bp bins and the number of reads per bin was normalized in counts per million (CPM) using the bamCoverage function of deeptools (https://deeptools.readthedocs.io/) (Ramirez *et al.*, 2016). Duke excluded regions list (wgEncodeDukeMapabilityRegionsExcludable table from UCSC genome browser) were removed. The resulting bigwig files were used for the study of chromosome Y coverage more specifically using deeptools multiBigwigSummary function. Chromosome Y was divided into 500 bp bins, and an average score was obtained for each of the samples in every bin. Regions with no reads in all samples were removed. To avoid chromosome X regions that are homologous to chromosome Y, bins that presented reads from samples corresponding to pregnancies carrying a female fetus were removed. Then, mean coverage of chromosome Y bins was calculated for those pregnancies of a male fetus.

6.3 DIFFERENTIALLY METHYLATED REGIONS

6.3.1 Detection

Differentially methylated regions (DMRs) were obtained using R-based software Methylaction (http://github.com/jeffbhasin/methylaction) (Bhasin *et al.*, 2016). The preprocessing of the bam files was performed using a window size of 50 bp and a fragment size of 370 bp, based on the library size for single-read sequencing data, and "paired" for paired-end sequencing data. Chromosome X and Y were excluded from the analysis.

DMR detection was performed using the methylaction() function with the following parameters: poifdr = 0.1, stageone.p = 0.05, anodev.p = 0.05, post.p = 0.05, freq = ½, minsize = 100, joindist = 100, adjust.var = NULL, nperms = 0, perm.boot = F.

6.3.2 Annotation

DMRs were annotated using R-based software Goldmine (https://github.com/jeffbhasin/goldmine) (Bhasin and Ting, 2016). The getGenes() function was used to obtain the ENSEMBL genes for hg19 for annotation. In cases of multiple overlaps, categories were established by using the priority: promoter, 3' end, exon, intron, intergenic. The getFeatures() function was used to obtain transcription factor binding sites (TFBS) feature set for annotation from the table 'wgEncodeRegTfbsClusteredV3'. This feature was concatenated with the output from getCpgFeatures() function, which gives annotation for CpG island context, resulting in one feature list with the genomic ranges each for ENCODE ChIP-seq peaks and CpG islands/shores/shelves (CpG shore, ±2 kb of islands; CpG shelf, ±2 kb of shores). DMRs were annotated for genomic context and for the above-mentioned features. The output corresponding to CpG island annotation was subsequently annotated for genomic context.

6.4 FUNCTIONAL ANALYSIS

For functional analysis, DMRs were filtered to keep those with p-value < 0.01 and |log2(fold change)| > log2(1). According to log2(fold change) values, regions were divided in hypermethylated and hypomethylated in sPE compared to control.

6.4.1 CpG Islands promoters

A list of genes with one or more Gene Ontology (GO) IDs associated was obtained from ENSEMBL using the BioMart data-mining tool (https://www.ensembl.org/biomart/martview/).

DMRs annotated as promoter that occurred at a CpG island were obtained. Genes annotated in these regions that were present in the list of genes with GO ID were subjected to functional analysis using Molecular Signatures Database v6.2 (MSigDB) with Gene Set Enrichment Analysis GSEA software (http://software.broadinstitute.org/gsea/msigdb/) (Subramanian *et al.*, 2005; Mootha *et al.*, 2003). Investigate gene set option was used to study enrichment for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways database. Top 10 gene sets with a false discovery rate (FDR) q-value below 0.05 were obtained.

6.4.2 Transcription factors binding sites

DMRs annotated in transcription factor binding sites (TFBS) in hyper- and hypomethylated regions were obtained. The number of times that each TFBS appeared in each group of DMRs was counted and divided by the total of DMRs annotated as TFBS in the group. Using ratio of appearance of each TFBS the DMRs were classified as overrepresented in sPE and overrepresented in control when the proportion of annotated TFBS differed by at least two-fold in hypermethylated compared to hypomethylated regions, and vice versa. A list of the transcription factors corresponding to the TFBS overrepresented in each group was subjected to functional analysis. Overrepresentation Enrichment Analysis (ORA) from WebGestalt (http://www.webgestalt.org) (Liao *et al.*, 2019) was used to study enriched biological process (http://www.geneontology.org). Genome protein coding was used as reference gene list, and top 10 enriched categories were obtained.

6.5 DMRS AND DISEASE

To avoid regions with shallow coverage, the list of DMRs was filtered to keep those in which $\geq$ 50% of samples of at least one group values are above the methylated cutoff point set during initial filtering in methylaction (http://github.com/jeffbhasin/methylaction) (Bhasin *et al.*, 2016). Of them, those annotated as promoter (distance to nearest gene = 0, promoter.per >

50) were kept. DMRs-associated genes were compared to BioMart list and those with associated GO IDs were used for the following analysis:

6.5.1 GAD enrichment

The list of genes annotated in the above-mentioned regions were subjected to disease ontology analysis. Functional annotation tool from DAVID online was used to study enrichment in the Genetic Association Database (GAD) (Becker *et al.*, 2004). Custom parameters were used (Count=10 and EASE=0.05) and top-5 significant enriched diseases were obtained.

6.5.2 Preeclampsia-related genes

Ingenuity Pathway Analysis (IPA) (QIAGEN) was used to obtain a list of genes related to "preeclampsia" and "maternal hypertension" terms. Moreover, a list of genes suspected to play a role in preeclampsia was obtained from Martin *et al.* (Martin *et al.*, 2015). Genes from both lists were joined and used as preeclampsia-related genes. DMRs' annotated genes were compared to the preeclampsia-related gene list, and those regions with annotated genes present in the list were kept.

6.6 CELL-FREE DNA FRAGMENTATION PATTERNS

The CollectInsertSizeMetrics function from PICARD software (https://broadinstitute.github.io/picard/index.html) was used to obtain insert size data from paired-end run bam files. Proportion of reads for different size ranges was obtained by dividing the number of reads with a size corresponding to the established range by the total of reads of the sample. Using our previously obtained fetal fraction estimation data, correlation between this variable and proportion of fragments was calculated.

6.7. STATISTICAL ANALYSIS

All analyses were performed in R (www.r-project.org) and GraphPad Prism (La Jolla, CA) statistical programs. R statistical programming language was used to test for normality of data (saphiro.test), to compare DMR annotation proportions between groups (prop.test) and to plot fragment distribution. R was also used to analyze correlation between fragment proportions and fetal fraction using ggpubr package. GraphPad Prism was used to plot

annotation distribution, and to plot and analyze data from chromosome Y quantification, as well as proportion of reads. A D'Agostino-Pearson omnibus test was used to test normality of the data; based on that result, either unpaired t-test or Mann- Whitney test was used. p-values ≤ 0.05 were considered statistically significant.

V. RESULTS

V. RESULTS

1. RUN 1

1.1 DIFFERENTIALLY METHYLATED REGIONS: ANNOTATION

The methylated fraction of cfDNA obtained by MBD-enrichment was sequenced, and regions with a significant difference in methylation between sPE and control group were obtained using Methylaction software. After filtering, 25019 DMRs with p-value < 0.01 and log2fc > |1| remained. Of them 13460 were hypermethylated in sPE, and 11559 were hypomethylated. To begin understanding the potential functions of DMRs, their genomic locus was examined by overlapping region coordinates with annotated genes of the reference genome (hg19) using Goldmine software. The percentage of DMRs of each type that annotated to each category is shown in Figure 10.

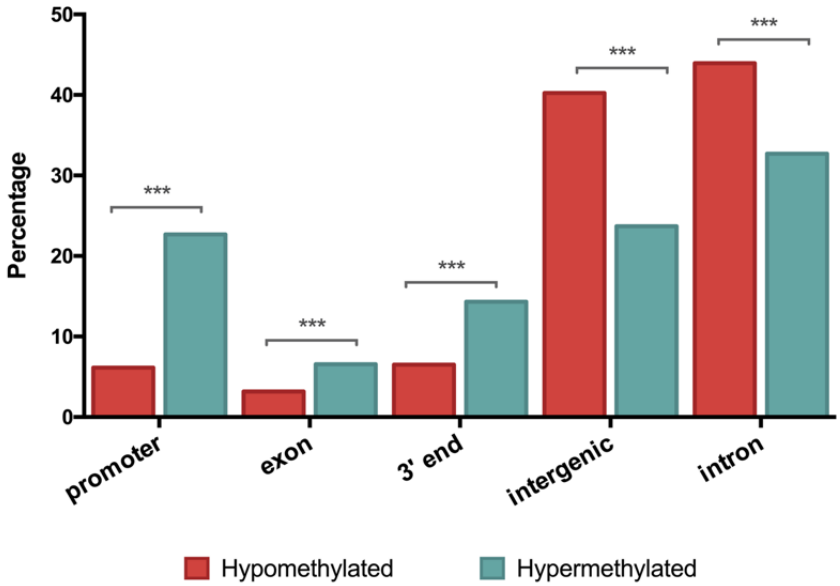

Figure 10. Genomic context for differentially methylated regions - Run 1. Percentage of hypomethylated (red) and hypermethylated (green) regions that overlap with the genomic context categories of intergenic, promoter, exon, intron and 3' end stop sites. *** p-value < 0.001

Annotation revealed that hypomethylated DMRs were enriched in intergenic (40.2%) and intron regions (43.9%), while hypermethylated regions were enriched in promoter (22.7%), intergenic (23.7%) and intron regions (32.7%). Comparison of proportions of annotated regions showed that all proportions were significantly different between types (p-value < 0.001). DMRs annotated as promoter, 3' end and exon regions were significantly higher in hypermethylated regions, while regions annotated as intergenic and intron were significantly

more represented in hypomethylated regions. The biggest difference between types of DMRs was in the percentage of DMRs annotated as promoter, with a 3.7-fold increase in hypermethylated compared to hypomethylated regions (6.1% in hypomethylated, 22.7% in hypermethylated).

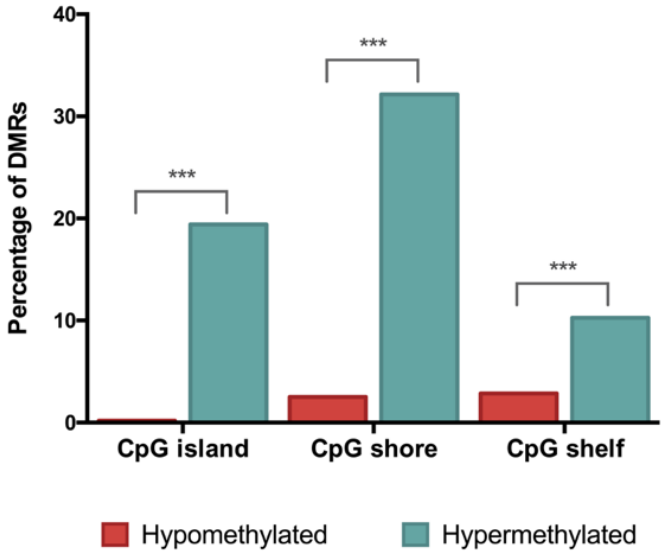

Figure 11. CpG island annotation - Run 1. Percentage of hypomethylated (red) and hypermethylated (green) regions that overlap with genomic feature sets of CpG island, CpG shore (±2 kb of islands) and CpG shelf (±2 kb of shores). *** p-value < 0.001

DMRs were also annotated with respect to CpG islands using the 'cpgIslandExt' table of the UCSC genome browser, which included islands, shores (±2 kb from these islands) and shelves (±2 kb from shores). From the total of DMRs of each type, 5.6% of hypomethylated and 61.8% of hypermethylated regions annotated to any of the CpG regions. For all CpG annotation categories, the proportion of DMRs corresponding to hypermethylated was significantly higher than hypomethylated regions (p-value < 0.001) (Figure 11). The greatest difference between the two types of DMRs occurred at CpG islands, where a 97.5-fold increase was observed (0.19% in hypomethylated, 19.4% in hypermethylated).

These observations prompted us to further investigate CpG islands. Regions annotated as CpG islands were assessed for genomic context.

Annotation showed that CpG islands hypermethylated in sPE were present in all the annotation categories, while in the hypomethylated regions, DMRs in CpG islands were only present in promoter, 3' end and intergenic regions (Figure 12A). In both types of DMRs, most

were annotated to promoter regions: 21 out of 23 hypomethylated regions and 1124 out of 2613 hypermethyated regions were CpG islands and promoter regions (Figure 12B).

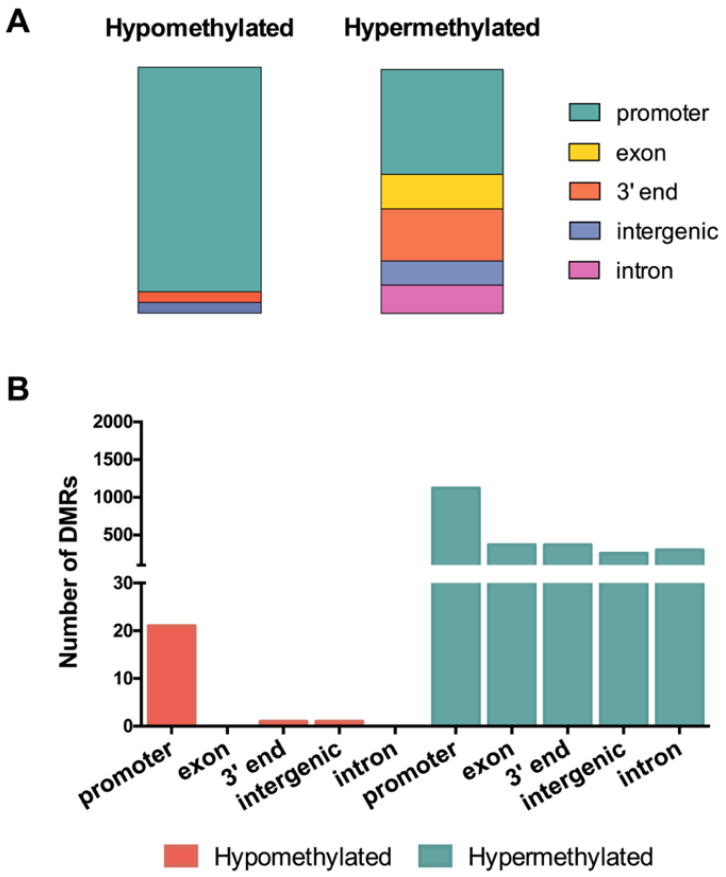

Figure 12. Genomic context for CpG island DMRs - Run 1. DMRs in CpG islands were annotated for genomic context. Proportion (**A**) and number (**B**) of DMRs of each type that overlap with annotation categories are shown.

Goldmine software was also used to examine TFBS present in the DMRs; a total of 35807 TFBS were found among 7419 regions. Of them, 33.1% were present in hypomethylated regions and 66.9% in hypermethylated regions. As one DMR can present many different TFBS, the ratio of TFBS per DMR was calculated in each group. In hypermethylated regions the mean number of TFBS per DMR was 5, while in hypomethylated regions it was 4.5. Nevertheless, most regions of both hyper- and hypomethylated regions had only one TFBS annotated (41% and 37.6%, respectively) (Figure 13).

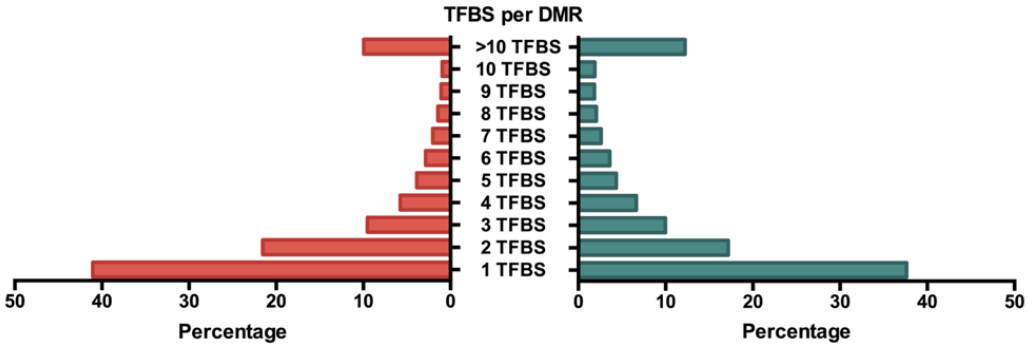

Figure 13. Number of transcription factor binding sites (TFBS) per DMR - Run 1. One DMR can contain more than one TFBS. Thus, the number of sites that overlap with one DMR were counted. Graph represents the percentage of hypomethylated (red) and hypermethylated (green) regions depending on the number of TFBS (1 to 10, more than 10) that overlap with one DMR.

1.2 FUNCTIONAL ANALYSIS

1.2.1 DMRs annotated as promoter and CpG island

With the aim of obtaining an overview of the role of methylation changes in severe preeclampsia, DMRs between sPE and Control corresponding to regions annotated as both promoter and CpG island (p-value < 0.01 and log2fc > |1|) were studied in detail. First, DMRs were divided into hypo- and hypermethylated regions and a list of the genes corresponding to the DMRs was obtained. To avoid pseudogenes, long non-coding RNAs (lncRNAs) and non-coding genes (ncRNA) that would add noise to functional analysis, DMR-associated genes were compared to a list of genes with an assigned GO ID. Shared gene were used for functional analysis.

Nine genes corresponding to hypomethylated regions, and 611 corresponding to hypermethylated regions were submitted to GSEA for over-representation analysis of KEGG pathways. No overlaps were found between genes from hypomethylated regions and KEGG pathways database. Figure 14 shows the top-10 pathways with statistical significance (FDR < 0.05) for the genes corresponding to hypermethylated promoters in sPE.

Among them, there are pathways related to signal transduction (MAPK signaling pathway, mTOR signaling pathway, VEGF signaling pathway, Calcium signaling pathway), endocrine system (Insulin signaling pathway, Adipocytokine signaling pathway), cell motility (Regulation of actin cytoskeleton) and cancer (Pathways in cancer, Melanoma, Acute myeloid leukemia).

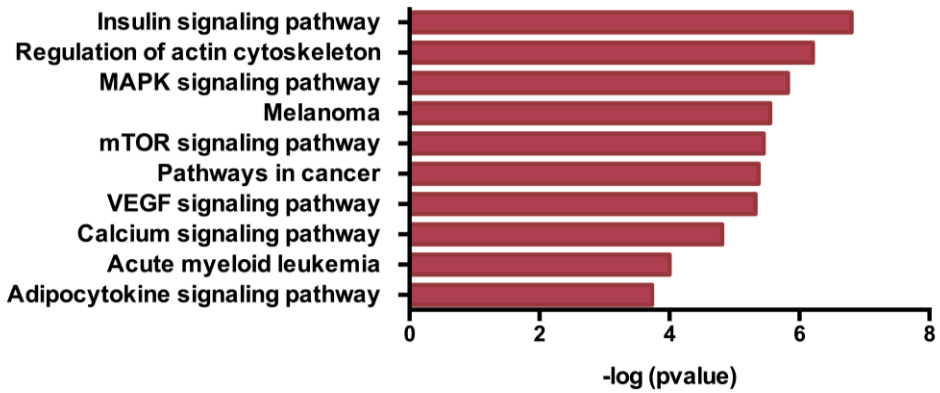

Figure 14. Functional analysis of CpG island promoters - Run 1. Regions that annotated to promoter regions that overlap with CpG islands were subjected to functional analysis using GSEA, and over-representation of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was obtained. Graph shows top-10 enriched KEEG pathways for hypermethylated CpG island promoters (FDR < 0.05). Y-axis represents enriched terms, significance is represented in x-axis as –log10(p-value). No significant enrichment was obtained for hypomethylated CpG island promoters.

1.2.2 DMRs annotated as TFBS

Because cfDNA in plasma is the result of nuclease fragmentation, it could be assumed that all fragments present in plasma were bound to proteins that protected them from nuclease degradation. Then, by studying cfDNA, bound proteins can be inferred. DMRs that have annotated TFBS were obtained, and those TFBS more than two-fold more represented in one type of DMRs than in the other were kept. As a result, 33 TFBS were classified as over-represented in hypomethylated regions, and 30 as over-represented in hypermethylated regions (Detailed in Annex 1). Transcription factors (TFs) corresponding to the TFBS were used to identify enriched biological processes. The top-10 most significant biological processes for each group (hypo- vs hypermethylated DMRs) are represented in Figure 15. Both types of DMRs showed one term in common (Negative regulation of cellular biosynthetic process); however, the other significantly enriched terms were distinct between types. Gene ontology showed that TFBS among hypomethylated regions corresponded to TFs enriched in processes implicated in positive regulation of transcription, gene expression and metabolic and biosynthetic processes (Figure 15A). In contrast, in the hypermethylated DMRs, the opposite pattern was observed: TFs corresponding to TFBS hypermethylated in sPE were significantly associated with negative regulation processes (Figure 15B).

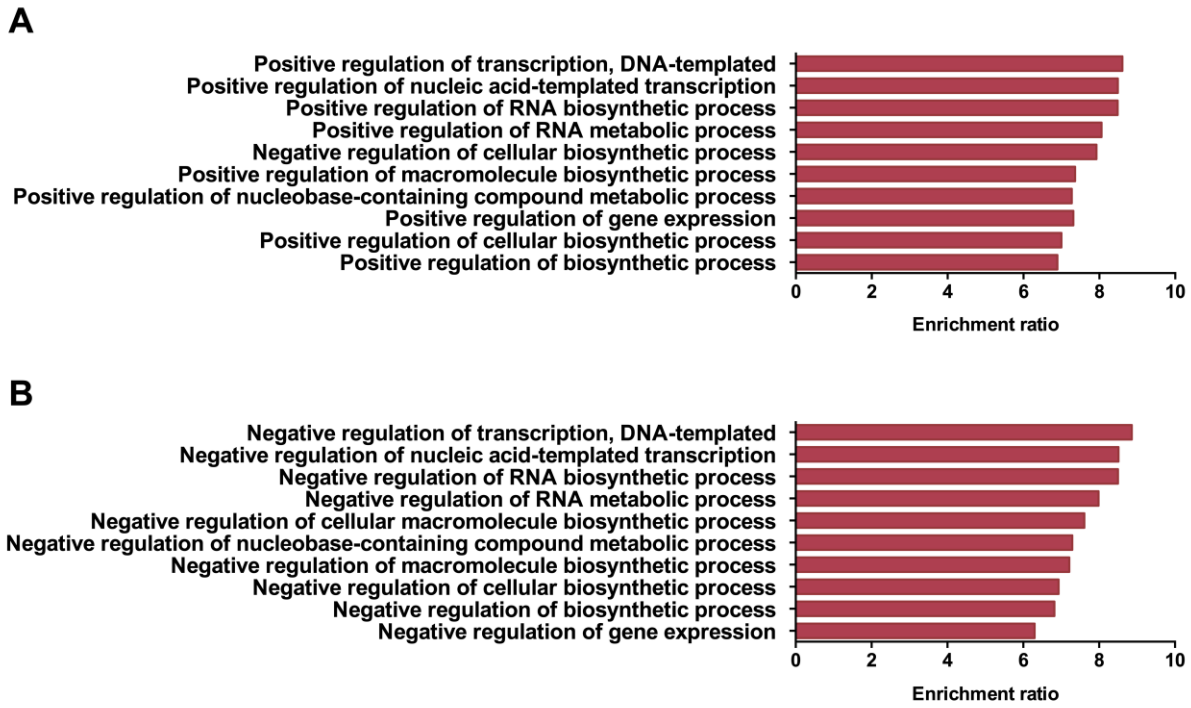

Figure 15. Functional analysis of TFBS - Run 1. Because one TFBS is present multiple times in the genome, the number of times each TFBS appeared in hypermethylated and hypomethylated regions were counted. TFBS were classified as overrepresented when the TFBS was more than two times in that group of DMR than in the other. Transcription factors corresponding to each group of overrepresented TFBS were subjected to Gene Ontology analysis using ORA from WebGestalt. Graph shows top-10 most significant biological processes associated to hypomethylated TFBS (**A**) and to hypermethylated TFBS (**B**). Y-axis represents enriched terms, x-axis represents enrichment ratio.

1.3. DMRS AND DISEASE

To explore the presence of relevant genes that could be related to preeclampsia physiopathology among DMRs, DAVID online tool was used to explore gene-disease association. Regions were filtered to obtain promoter regions with good coverage in the samples, and DMR-associated genes were obtained. As a result, 1235 genes corresponding to regions hypermethylated in sPE were used to test gene-disease association using Gene Association Database. The 5 most significant terms are shown in Figure 16. Of the list of genes, 187 were associated with type 2 diabetes, 68 with hypertension, 36 with apoplexy and cerebral hemorrhage, 28 with high-density lipoprotein (HDL) cholesterol levels, and 18 with kidney diseases.

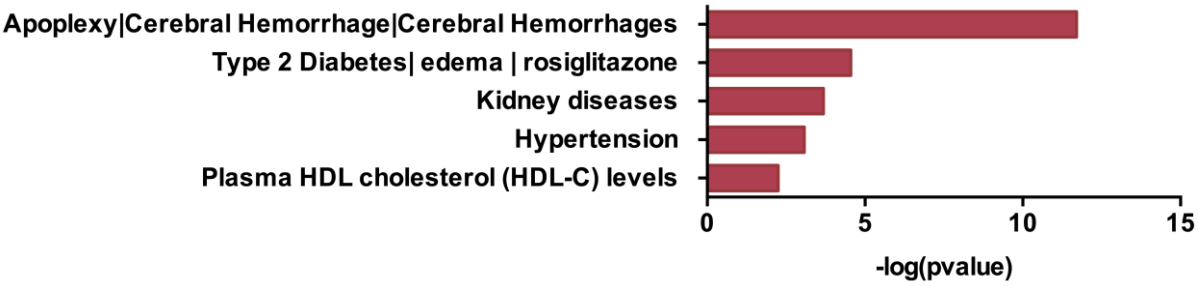

Figure 16. Gene-association analysis - Run 1. DMR-associated genes were obtained from good coverage promoter regions. No hypomethylated region met the requirement. DAVID online tool was used to explore gene-disease association among hypermethylated. Graph shows top 5 significant terms (y-axis). Significance is represented as -log10(p-value) in x-axis.

1.4. FETAL CELL-FREE DNA

To estimate fetal contribution to cell-free DNA, chromosome Y data were used to estimate fetal fraction. Sequencing data from all samples was used to remove regions corresponding to chromosome Y regions with high homology to chromosome X, which would lead to an overestimation of fetal fraction. Normalized coverage of chromosome Y filtered regions was calculated for each sample corresponding to pregnancies carrying a male fetus (Control n=5, sPE n= 5). Mean coverage of regions was calculated for each sample, and data were compared between groups (Figure 17). No significant differences were observed (Mean ± SEM: Control = 0.0085 ± 0.002, sPE= 0.007 ± 0.001; p-value = 0.6667).

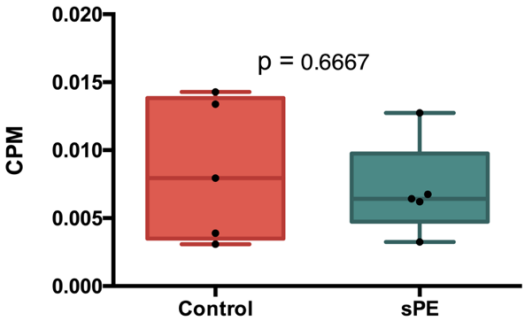

Figure 17. Fetal contribution to cfDNA – Run 1. Chromosome Y data were used to estimate fetal fraction, and only pregnancies carrying a male fetus were included in the analysis (Control n=5, sPE n=5). Box-and-whisker charts represent normalized mean coverage for chromosome Y regions in control (red) and sPE (green) samples. CPM= Counts Per Million, p-value= 0.667

2. RUN 2

2.1 DIFFERENTIALLY METHYLATED REGIONS: ANNOTATION

Following the same procedure as for the single-read run, sequences obtained from paired-end sequencing of the methylated fraction of cfDNA were used to performed DMR detection using Methylaction software. The resulting DMRs were filtered using p-value < 0.01 and log2fc > |1|, and a total of 134873 regions were used for the following analysis. First, DMRs were divided into hypermethylated (22042 DMRs) and hypomethylated (112831 DMRs) in sPE compared to control. Then, DMRs were annotated using Goldmine software as a first step in the understanding of their potential function. In Figure 18, the percentage of DMRs of each type that annotated to each category is shown.

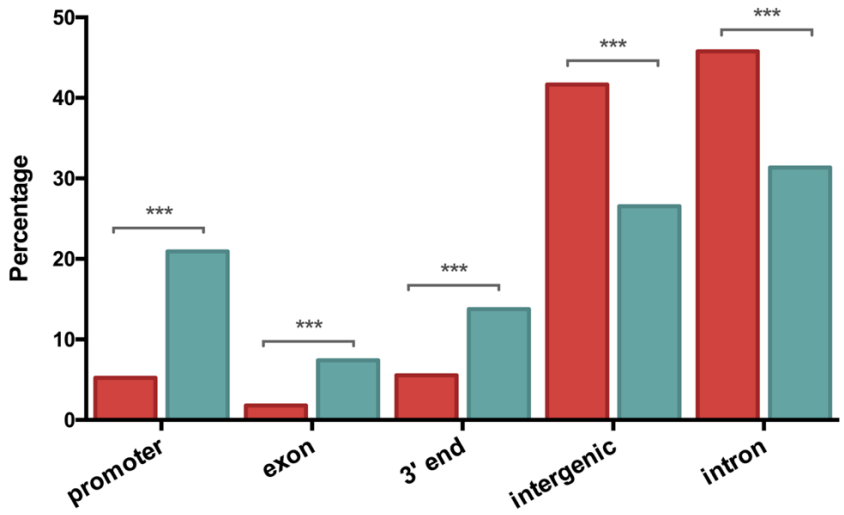

Figure 18. Genomic context for DMRs - Run 2. Percentage of hypomethylated (red) and hypermethylated (green) regions that overlap with the genomic context categories of intergenic, promoter, exon, intron and 3' end stop sites. *** p-value < 0.001

Annotation of DMRs based on genomic context illustrated that most hypomethylated DMRs corresponded to intergenic (41.7%) and intron regions (45.8%). In hypermethylated regions, apart from intergenic (26.5%) and intron (31.4%), DMRs were also enriched in promoter regions (20.9%). Comparison of the distribution of DMRs between groups revealed that the proportion of intron and intergenic DMRs were significantly higher in hypomethylated regions, while promoter, 3' end and exon regions were significantly more represented in hypermethylated DMRs (p-value < 0.001). The greatest difference corresponded to exon regions (4.2-fold increase: 1.7% in hypomethylated, 7.4% in hypermethylated), followed by a 4-fold increase in promoter regions (5.2% in hypomethylated, 20.9% in hypermethylated).

The distribution of DMRs with respect to CpG island context was also examined. From the total of DMRs corresponding to each type, 5.83% of hypomethylated and 58.3% of hypermethylated regions annotated to CpG islands, CpG shores or CpG shelves (Figure 19).

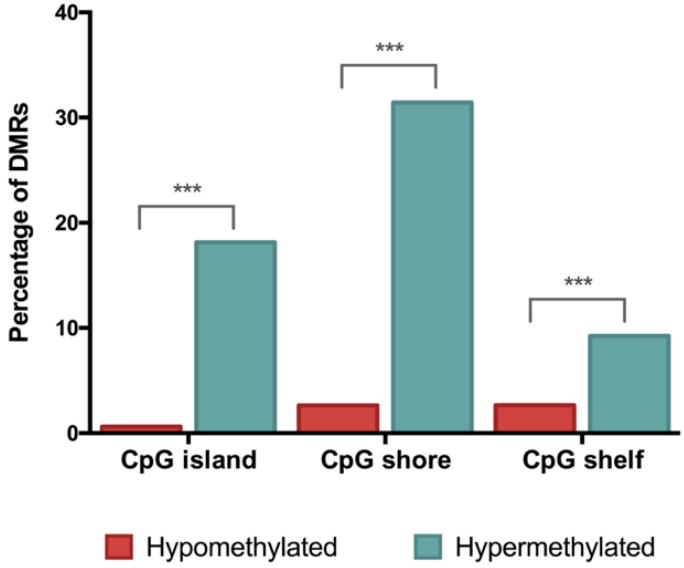

Figure 19. CpG island annotation - Run 2. Percentage of hypomethylated (red) and hypermethylated (green) regions that overlap with genomic feature sets of CpG island, CpG shore (±2 kb of islands) and CpG shelf (±2 kb of shores). *** p-value < 0.001

Comparison of the percentage of DMRs in the three different categories showed that the proportion of DMRs in all categories was significantly higher in hypermethylated regions (p-value < 0.001). The biggest difference corresponded to CpG islands, in which a 30.9-fold increase was observed in the percentage of hypermethylated compared to hypomethylated regions. Considering that the total number of DMRs in each type differs, it is important to highlight that there is not only an increase in the percentage of DMRs but also in the total number of DMRs annotated as CpG islands, shores or shelves (6588 hypomethylated, 12968 hypermethylated).

Then, CpG island DMRs were annotated regarding genomic context to obtain their distribution in the genome (Figure 20). CpG islands from both types of DMRs were present in all the genomic context categories (Figure 20A), with promoter regions being the most represented in each (83.2% hypomethylated, 37.8% hypermethylated). Although the percentage of CpG islands in promoter regions is higher in hypomethylated regions, the total number of DMRs

annotated as CpG island and promoter regions was higher in the regions hypermethylated in sPE (553 hypomethylated, 1512 hypermethylated) (Figure 20B).

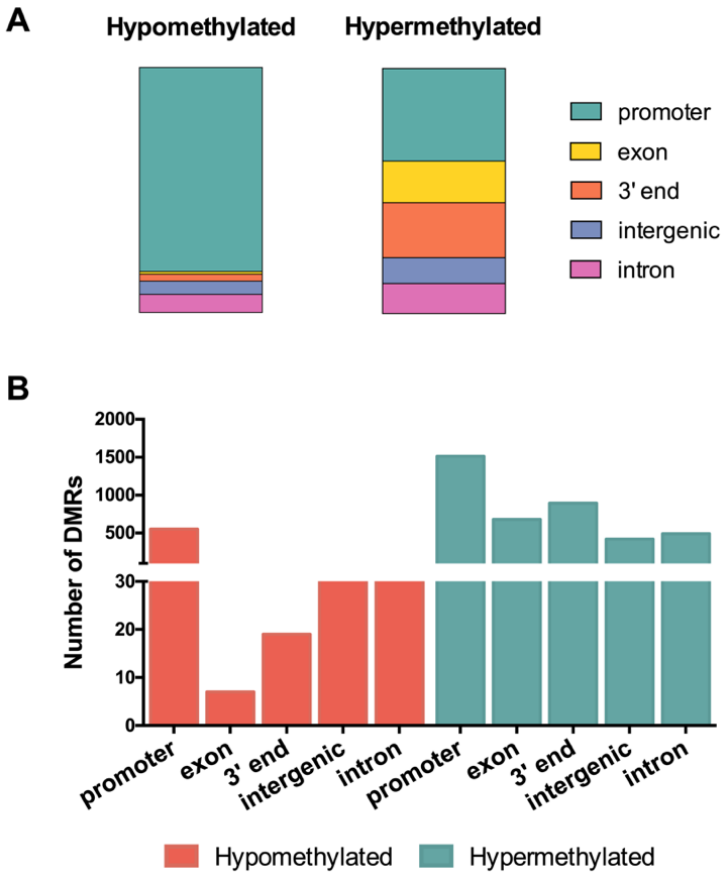

Figure 20. Genomic context for CpG island DMRs - Run 2. DMRs in CpG islands were annotated for genomic context. Proportion (**A**) and number (**B**) of DMRs of each type that overlap with annotation categories are shown.

Annotation of DMRs for TFBS by Goldmine showed that 149842 TFBS were present in 30258 DMRs. Of these sites, 21.5% were in hypermethylated regions and 78.5% in hypomethylated regions. The mean number of TFBS that fell within one DMR was 4.5 in hypermethylated regions and 5.1 in hypomethylated regions, but for both types, most of the DMRs only had one TFBS (41.8% of hypomethylated, 41.4% of hypermethylated) (Figure 21).

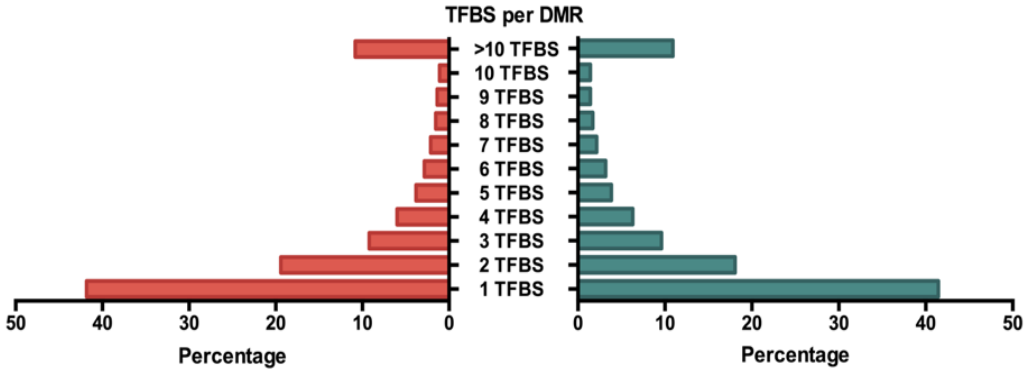

Figure 21. Number of TFBS per DMR - Run 2. Graph represents the percentage of hypomethylated (red) and hypermethylated (green) regions depending on the number of TFBS (1 to 10, more than 10) that overlap with one DMR.

2.2 FUNCTIONAL ANALYSIS

2.2.1 DMRs annotated as promoter and CpG island

DMRs annotated as promoter that occurred at a CpG island were compared to a Biomart list of genes with an associated ID, to remove those that were non-protein-coding genes. From the total of genes, 744 hypermethylated associated genes and 297 hypomethylated associated genes were subjected to functional analysis. Over-representation analysis of KEGG gene sets was carried out using MSigDB from GSEA software. Top-10 enriched pathways for each type of DMR-associated genes is shown in Figure 22.

Among the enriched pathways associated to hypomethylated genes (Figure 22A), we found signal transduction pathways (Wnt signaling pathway, MAPK signaling pathway), metabolism pathways (N-Glycan biosynthesis, Selenoamino acid metabolism) and immune system pathways (Leukocyte transendothelial migration). Genes annotated in hypermethylated regions (Figure 22B) were significantly enriched in signal transduction pathways (Calcium signaling pathway, mTOR signaling pathway, VEGF signaling pathway, Notch signaling pathway). Moreover, there are pathways related to endocrine (Insulin signaling pathway) and circulatory systems (Vascular smooth muscle contraction).

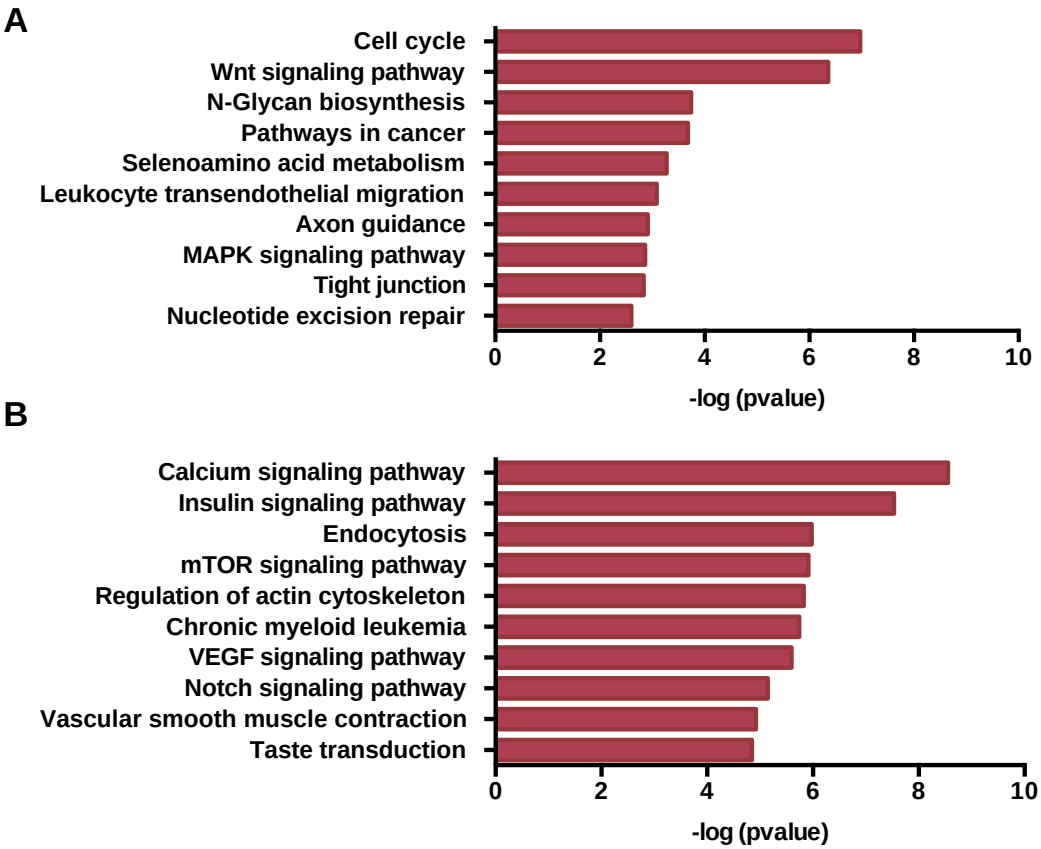

Figure 22. Functional analysis of CpG island promoters - Run 2. Regions that annotated to promoter regions that overlap with CpG islands were subjected to functional analysis using GSEA, and over-representation of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was obtained. Graph shows top-10 enriched KEEG pathways for hypomethylated (**A**) and hypermethylated (**B**) CpG island promoters (FDR < 0.05). Y-axis represents enriched terms, significance is represented in x-axis as –log10(p-value).

2.2.2 DMRs annotated as TFBS

Regions with annotated TFBS for hyper- and hypomethylated DMRs were used to perform Gene Ontology analysis. First, both types of DMRs were compared, and 51 TFBS and 15 TFBS appeared to be over-represented among hypo- and hypermethylated regions, respectively (Detailed in Annex 2). Human gene ontology term profiling of the TFs corresponding to the TFBS was performed using WebGestalt. Top-10 enriched terms were obtained for each DMR type and are represented in Figure 23.

Three terms were shared between both types of DMR (Positive regulation of nucleobase-containing compound metabolic process, Positive regulation of RNA metabolic process and Negative regulation of gene expression). Among the enriched processes corresponding to hypomethylated regions (Figure 23A), multiple terms were accompanied by their opposite in the list of enriched terms: for regulation of biosynthetic process, regulation of gene expression

and regulation of cellular biosynthetic process, both positive and negative regulation were associated with the TFs. The others were terms associated with positive regulation of biological processes. On the other hand, for the group of TFs corresponding to hypermethylated regions in sPE (Figure 23B), TFBS were mostly associated with negative regulation of biological processes. In this case, both positive and negative regulation of transcription by RNA pol II was present among the enriched terms.

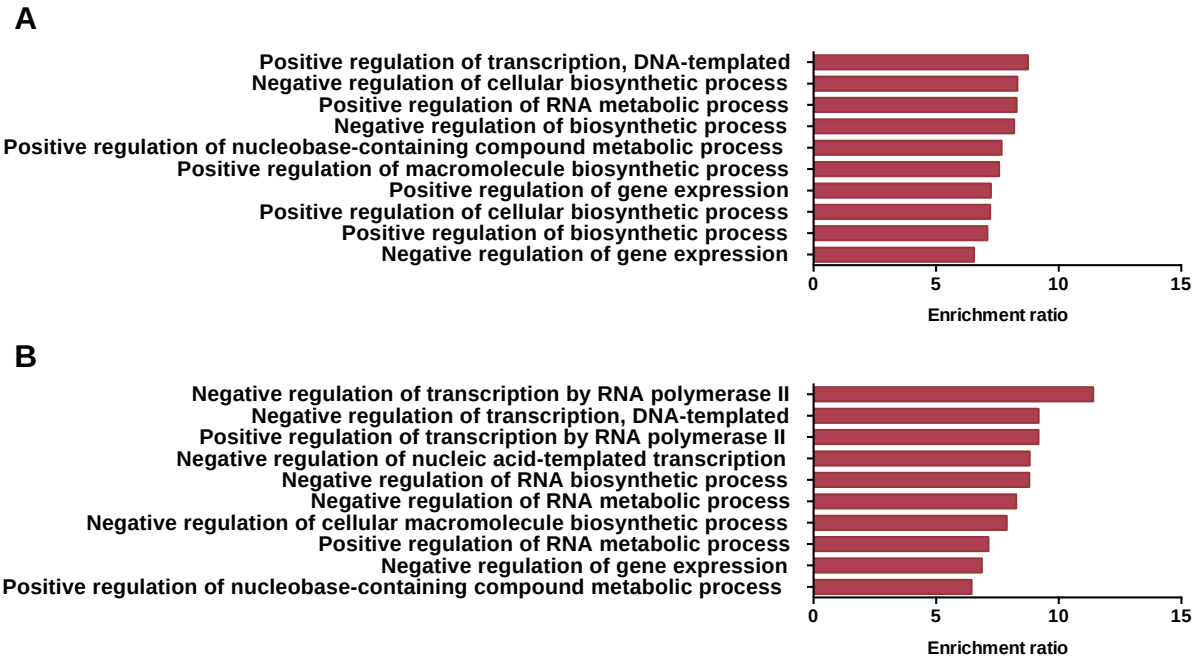

Figure 23. Functional analysis of TFBS - Run 2. Due to the fact that one TFBS is present multiple times in the genome, number of times of each TFBS appeared in hypermethylated and hypomethylated regions were counted. TFBS were classified as overrepresented when the TFBS was more than two times in that group of DMRs than in the other. Transcription factors corresponding to each group of overrepresented TFBS were subjected to Gene Ontology analysis using ORA from WebGestalt. Graph shows top-10 most significant biological processes associated to hypomethylated TFBS (**A**) and to hypermethylated TFBS (**B**). Y-axis represents enriched terms, x-axis represents enrichment ratio.

2.3. DMRS AND DISEASE

Gene-disease association was used to test the enrichment of DMR-associated genes with the aim of investigating the role of methylation changes in preeclampsia symptoms. Promoter DMRs were filtered to avoid regions with shallow coverage, and the list of resulting DMR-associated genes was obtained. In total, 2441 genes corresponding to the regions hypermethylated in sPE were used to explore gene-disease association. Top-5 significant terms are shown in Figure 24.

Of the list of genes, 340 were associated with type 2 diabetes, 116 with hypertension, 47 with HDL cholesterol levels, 46 with apoplexy and cerebral hemorrhage and 14 with coronary spastic angina (Figure 24).

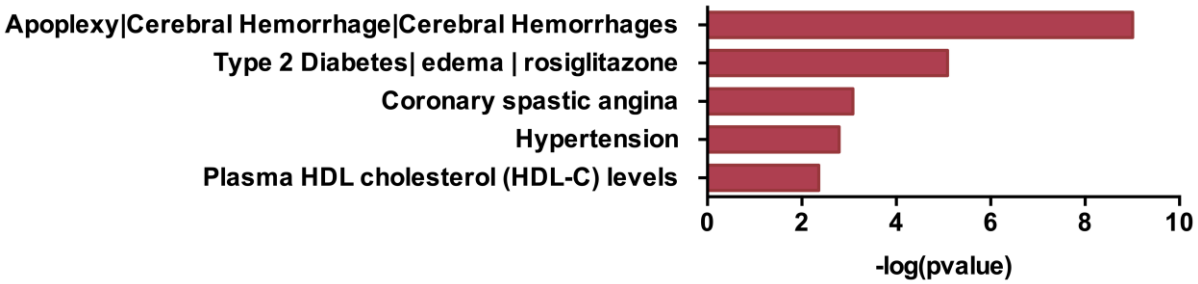

Figure 24. Gene-association analysis - Run 2. DMR-associated genes were obtained from good coverage promoter regions. No hypomethylated region met the requirement. DAVID online tool was used to explore gene-disease association among hypermethylated. Graph shows top-5 significant terms (y-axis). Significance is represented as -log10(p-value) in x-axis.

Due to the high number of genes associated with diseases, a more targeted approach was used to explore DMR-associated genes list with regard to genes that can be involved in the mechanisms underlying preeclampsia. The filtered DMR list was compared to genes previously related with preeclampsia obtained from IPA database and publications (Martin *et al.*, 2015). A total of 147 DMRs occurred at promoters of genes related to the disease; these DMRs corresponded to 88 different genes such as *ENG, MMP23B, DNMT3A* (Detailed in Annex 3). Among them we found 3 DMRs overlapping the *eNOS* gene (also known as *NOS3*) (Figure 25).

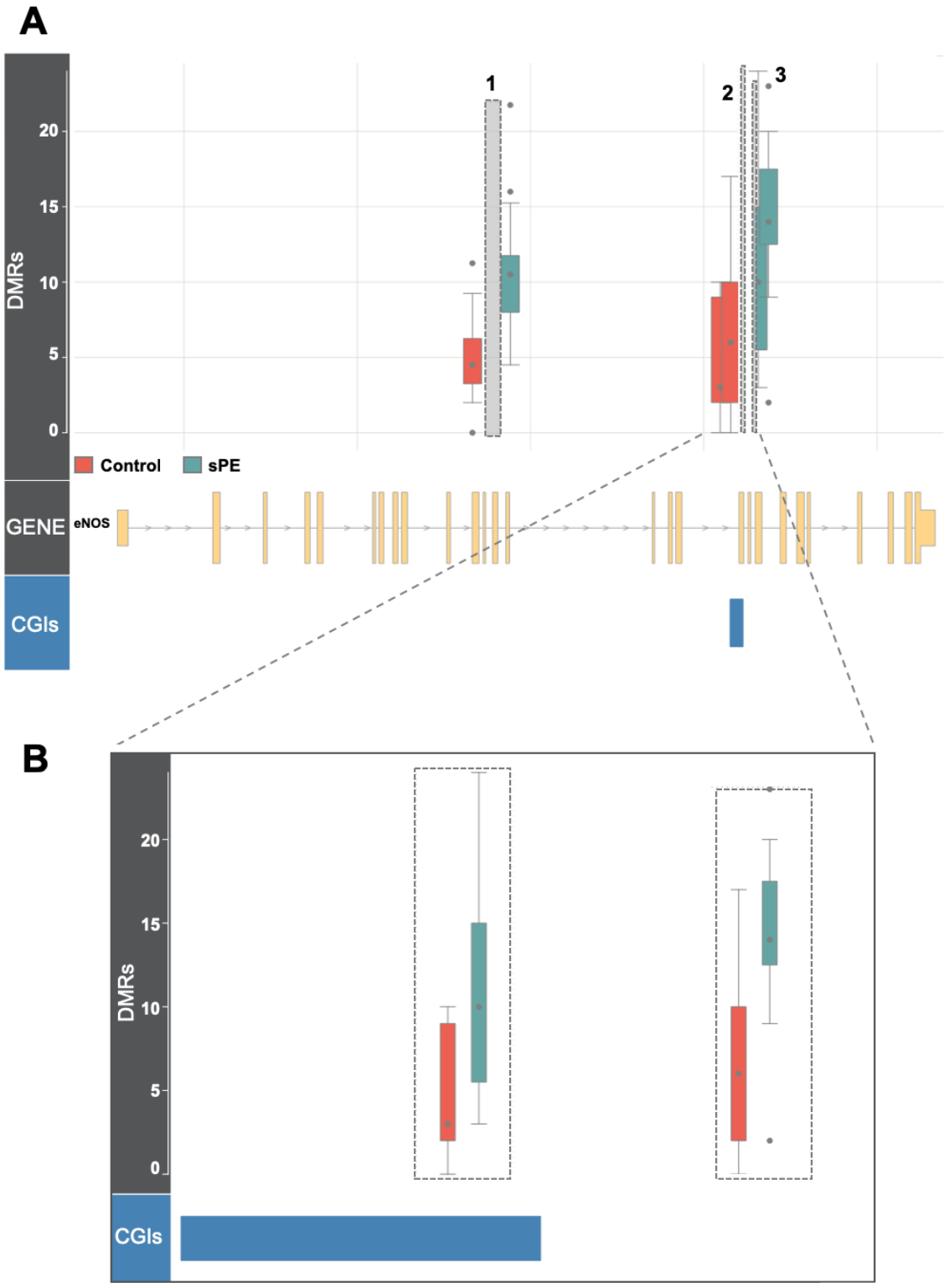

Figure 25. DMRs in the *eNOS* gene. The 3 DMRs found in *eNOS* are represented (**A**). DMR panel shows boxplots with normalized read depth values (x-axis) of Control (red, n= 22) and sPE (green, n=19) samples for each DMR (dashed grey boxes). Gene panel shows *eNOS* gene, and CpG islands present in the region are represented in CGI panel as a blue rectangle. (**B**) Detail of region containing DMRs 2 and 3, represented as dashed boxes. *eNOS= Endothelial Nitric Oxide Synthase*

2.4. FETAL CELL-FREE DNA

Estimation of fetal contribution to cell-free DNA was performed as in Run 1. Normalized coverage of regions annotated to selected chromosome Y regions was obtained for each sample corresponding to pregnancies carrying a male fetus (Control n= 11, sPE n= 8). Mean coverage of regions was calculated for each sample, and data were compared between groups (Figure 26). Although the fetal fraction was higher in the Control group, no significant differences were observed between groups (Mean ± SEM: Control = 0.002 ± 0.0007, sPE= 0.001 ± 0.0003; p-value = 0.1179).

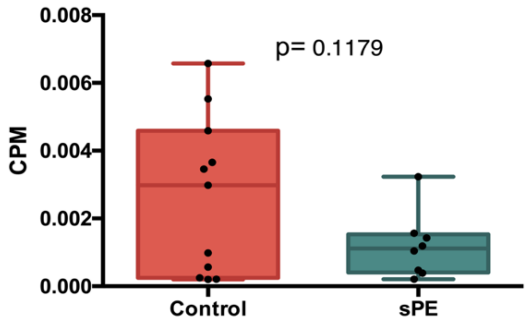

Figure 26. Fetal contribution to cfDNA- Run 2. Chromosome Y data were used to estimate fetal fraction, then only pregnancies carrying a male fetus were included in the analysis (Control n=11, sPE n=8). Box and whisker charts represent normalized mean coverage for chromosome Y regions in control (red) and sPE (green) samples. CPM= Counts Per Million, p-value= 0.1179

2.5. SIZE DISTRIBUTION OF CELL-FREE DNA IN PLASMA SAMPLES

In paired-end sequencing, size of each sequenced cfDNA molecule can be deduced from the positions of the reads. Called “insert size,” this corresponds to the size of the molecule between sequencing adapters. Here, insert size distribution was obtained for each sample using picard software. Size distribution for each sample is shown in Figure 27.

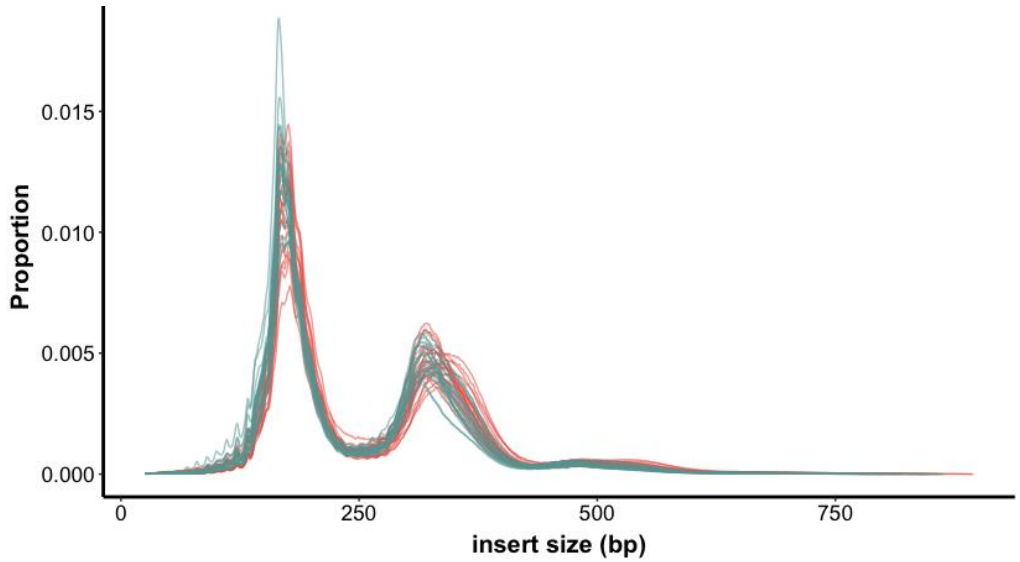

Figure 27. Fragment distribution of cell-free DNA molecules. Graph shows fragment size distribution of 41 samples from 19 (green) pregnant women with sPE diagnosis and 21 (red) control pregnancies. Molecule sizes and proportions are represented on the x-axis and y-axis, respectively.

Mean size distributions of cfDNA fragments for each group are shown in Figure 28A. Irrespective of the group, fragments showed the expected pattern corresponding to linear progression of nucleosome units wi th peaks occurring at 10-bp periodicity at sizes shorter than 150 bp (Figure 28B).

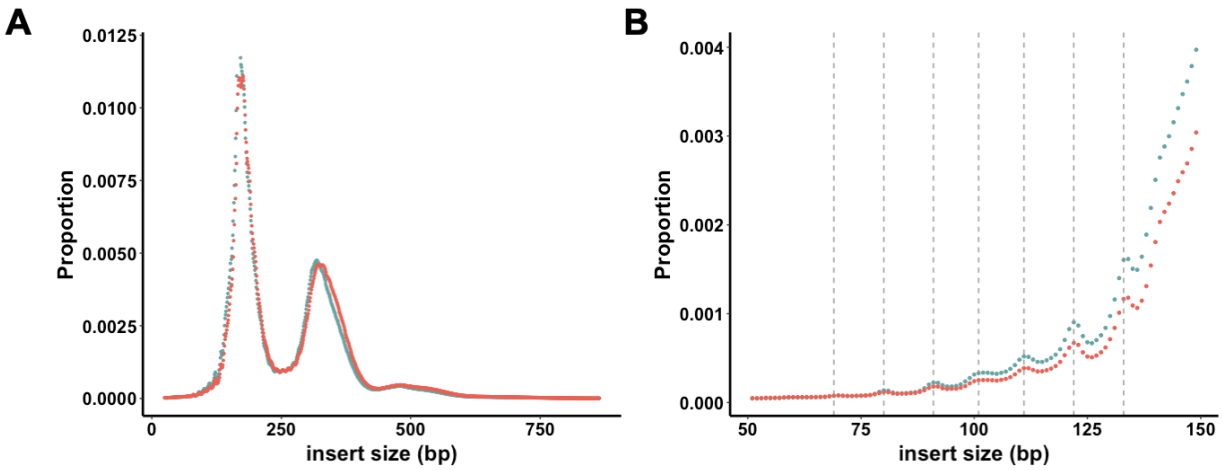

Figure 28. Mean fragment distribution of cell-free DNA molecules. Data from paired-end sequencing was used to obtain size of molecules present in cell-free DNA (cfDNA). **A:** Mean size distribution of cfDNA fragments in sPE (green) and Control (red) groups are shown; x-axis represents molecules size and y-axis represents proportion of molecules. **B**: Detail of proportion of molecules between 50 and 150 bp. Mean proportion is represented for sPE (green) and control (red) samples. Vertical dashed lines indicate peaks with ~ 10-bp periodicity.

To quantify the observed differences in size distribution between groups, proportion of DNA fragments was calculated for selected regions based on observed distribution differences.

First, a range between 20 and 150 bp was established to measure “short fragments” (Yu *et al.*, 2014). For that, proportion of reads with sizes within the range were quantified in each sample and compared between groups (Figure 29).

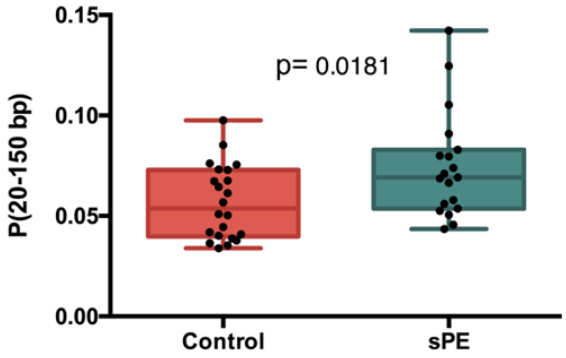

Figure 29. Proportion of molecules between 20 and 150 bp. Box-and-whiskers charts represent proportion of molecules with a size between 20 and 150 bp for Control (red) and sPE (green) groups. p-value = 0.0181

Methylated cfDNA fragments were significantly enriched in short fragments in the sPE group compared to Control group (Mean ± SEM: Control=0.056 ± 0.003, sPE= 0.0744 ± 0.006; p-value= 0.0181) (Figure 29).

In view of the differences observed in size distribution (Figure 28) in the proportions of fragments corresponding to the position of the mono- and di-nucleosome sizes, proportion of reads was measured in the ranges 120-170 bp, 175-225 bp, 270-320 bp and 325-470 bp.

Significant differences were found between groups for all size ranges. Proportion of fragments between 120 and 170 bp was significantly higher in sPE than in control group (Mean ± SEM, Control= 0.1862 ± 0.007, sPE= 0.2296 ± 0.011, p-value = 0.004) (Figure 30A). In contrast, in the range of 175-225 bp, fragment proportion was significantly higher in control samples (Control= 0.2667 ± 0.007, sPE = 0.2494 ± 0.003, p-value= 0.0425) (Figure 30B). Similarly, in the second peak, proportion of 270-320 bp fragments was significantly higher in the sPE group (Control= 0.1290 ± 0.005, sPE=0.1454 ± 0.004, p-value= 0.0217) (Figure 30C) and the opposite pattern was observed in the proportion of 325-470 bp fragments (Control= 0.2514 ± 0.008, sPE= 0.2153 ± 0.009, p-value= 0.0073)(Figure 30D).

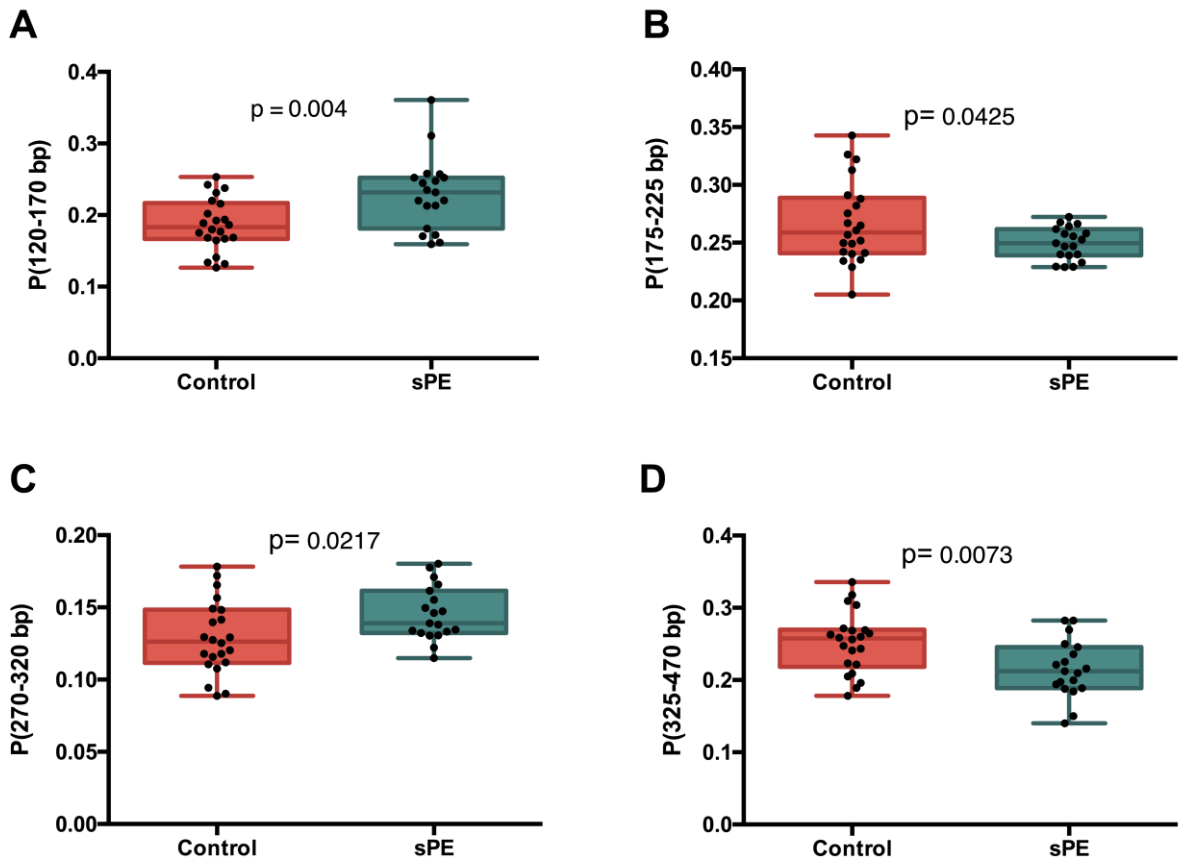

Figure 30. Cell-free DNA molecules fragment distribution differences. Box-and-whiskers charts represent proportion of molecules with a size between 120 and 170 bp (**A**, p-value =0.004), between 175 and 225 bp (**B**, p-value= 0.0425), between 270 and 320 bp (**C**, p-value= 0.0217) and between 325 and 470 bp (**D**, p-value 0.0073) for Control (red) and sPE (green) groups.

2.6. SIZE DISTRIBUTION AND FETAL CONTRIBUTION

As cfDNA from fetal origin is generally shorter than maternal DNA (Yu *et al.*, 2014), the relationship between fetal fraction and fragment distribution was assessed. Following this idea, the relationship between short fragments and fetal contribution was studied using fetal fraction data for pregnancies carrying a male fetus. First, the relationship between 20-150 bp fragments and fetal contribution was assessed, and no significant correlation was observed between both variables (R=-0.33, p=0.17) (Figure 31).

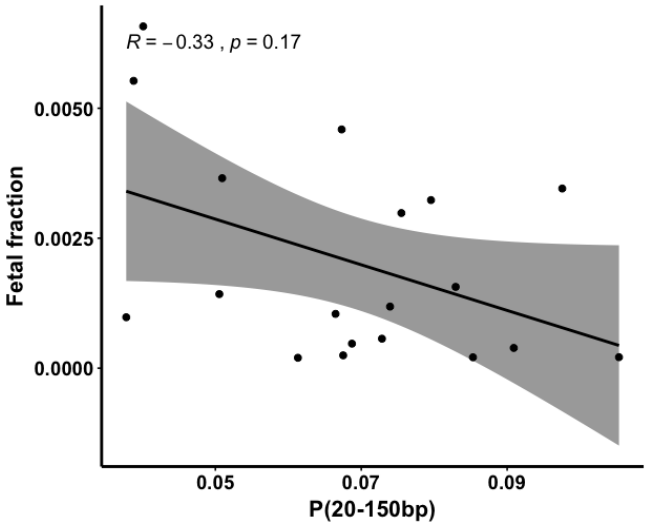

Figure 31. Correlation between cell-free DNA short fragments and fetal fraction. Fetal fraction estimation from pregnancies carrying a male fetus (Control n=11, sPE n=8) was used to evaluate its association with quantity of short fragments in cfDNA. Scatter plot shows relationship between fetal fraction (y-axis) and proportion of fragments between 20 and 150 bp (x-axis) irrespective of their group of origin. R =-0.33, p-value=0.17. Linear regression line is shown. *R= Spearman's correlation coefficient*

To find fragment proportions that might be related to fetal fraction using previously obtained data, the two samples with the highest (n=2, Control) and lowest fetal fraction (n=2, Control) were selected and fragments distribution was analyzed (Figure 32).

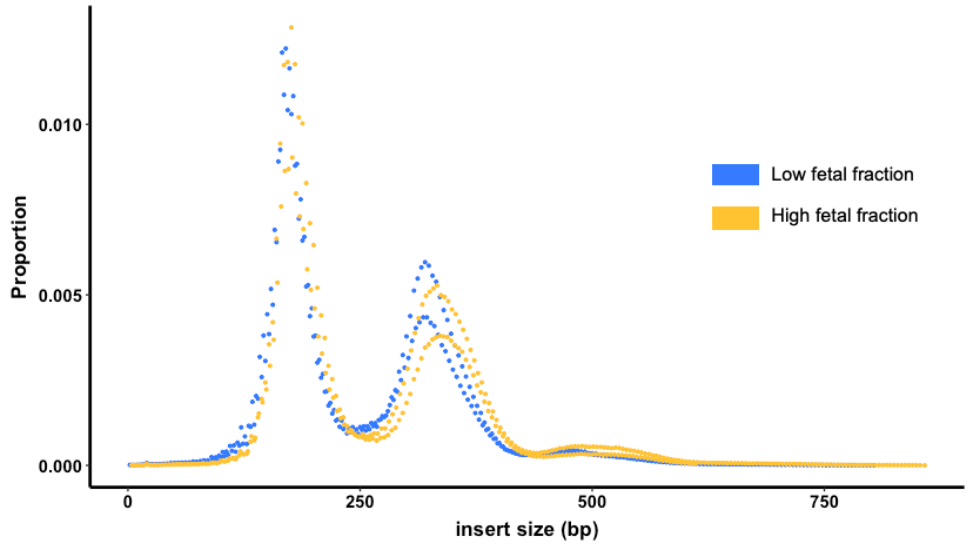

Figure 32. Fragment distribution in low and high fetal fraction samples. Using fetal fraction estimation data from pregnancies carrying a male fetus from both Control pregnancies and women with sPE, samples with high and low fetal fraction were selected. Plot shows proportion of fragments (y-axis) depending on their size (x-axis) for the 2 samples with the highest fetal fraction (yellow), and the 2 samples with the lowest fetal fraction (blue).

Differences between low and high fetal fraction samples were around the mono- and di-nucleosome peaks. The most obvious difference occurred at the 2-nucleosome peak: peak maximum was around 320 bp in the high fetal fraction samples, but smaller than expected in the low fetal fraction samples (~300 bp) (Figure 32).

Similarly, in the 1-nucleosome peak, high fetal fraction samples presented fewer fragments in the rising part of the peak than low fetal fraction samples, and the opposite was observed in the failing part of the peak. Then, proportion of fragments for 120-170 bp, 175-225 bp, 270-320 bp and 325-470 bp for samples corresponding to male fetus pregnancies were used to study correlation with fetal fraction (Figure 33).

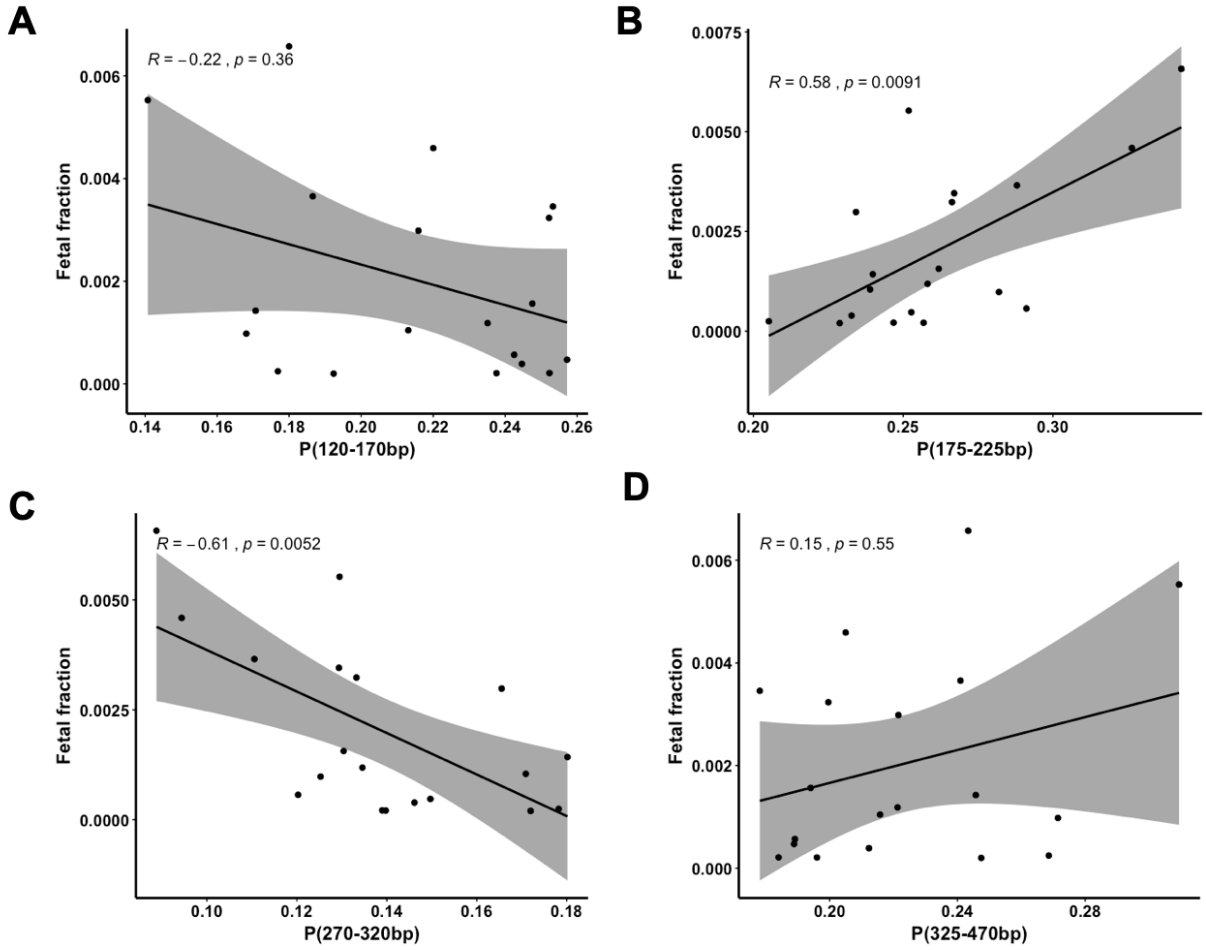

Figure 33. Association between cell-free DNA fragment size and fetal fraction. Fetal fraction estimation data from pregnancies carrying a male fetus (Control n=11, sPE n=8) were used to evaluate a relationship with cfDNA molecule size. Scatter plots show association between fetal fraction (y-axis) and proportion of cfDNA fragments in different size ranges (x-axis): **A**: 120-170 bp (R =-0.22, p-value= 0.36), **B**: 175-225 bp (R=0.58, p-value= 0.0091), **C**: 270-320 bp (R=-0.61, p-value= 0.0052), **D**: 325-470 bp (R=0.15, p-value=0.55). Linear regression line is shown in all the plots. *R= Spearman's correlation coefficient*

No correlation was found for proportion of fragments of 120-170 bp (R= -0.22, p-value= 0.36, Figure 33A) or 325-470 bp (R= 0.15, p-value= 0.55, Figure 33D) with fetal fraction. On the other hand, a moderate positive correlation was observed between fetal fraction and proportion of 175-225 bp fragments (R= 0.58, p-value= 0.0091, Figure 33B) and a moderate negative

correlation was observed between fetal fraction and proportion of 270-320 bp fragments (R= -0.61, p-value= 0.0052, Figure 33C).

Both size ranges corresponded to ranges with significant differences between sPE and Control groups (Figure 30B-C). Since the method used for fetal fraction estimation is based on chromosome Y, the effect of fetal fraction on the observed differences between groups could not be completely assessed. Data distribution showed a certain bias related to sample origin (Figure 34).

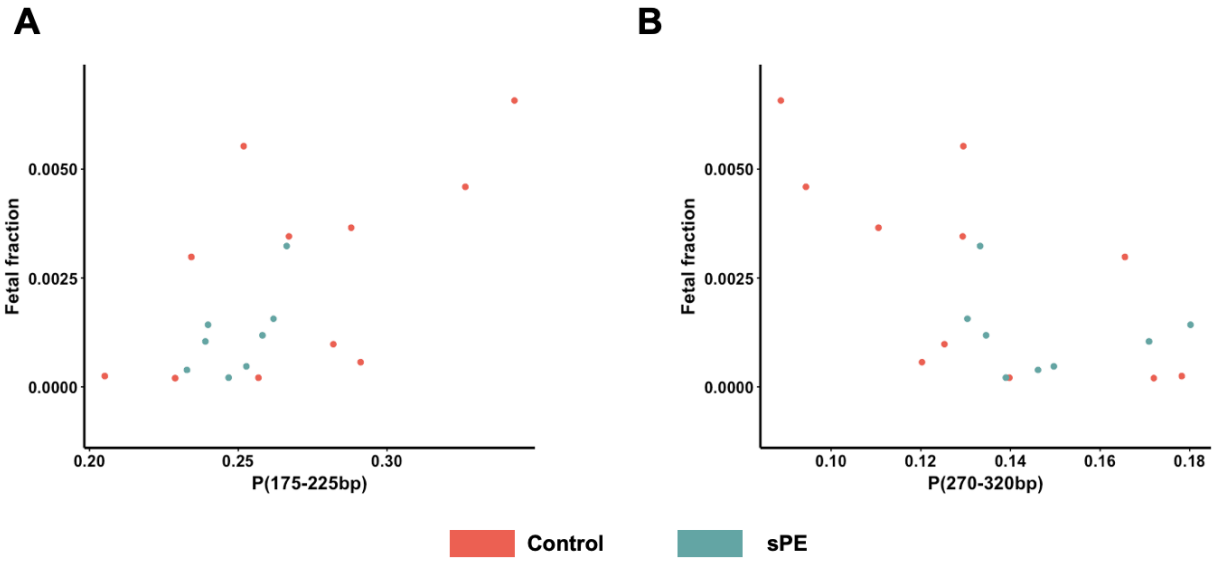

Figure 34. Effect of group of origin of samples on the observed distribution of fragment size. Scatter plots show sample distributions depending on their fetal fraction (y-axis) and proportion of cfDNA molecules (x-axis) between 175 and 225 bp (**A**) and 270 and 320 bp (**B**). Only pregnancies carrying a male fetus are included in the analysis. Green dots correspond to sPE samples (n=8) and red dots correspond to control samples (n=11).

Analysis on the proportion of fragments corresponding to 320 bp did not reveal any significant difference between sPE and control (Figure 35B). Correlation test revealed a moderate negative correlation between proportion of 320 bp fragments and fetal fraction (Figure 35A).

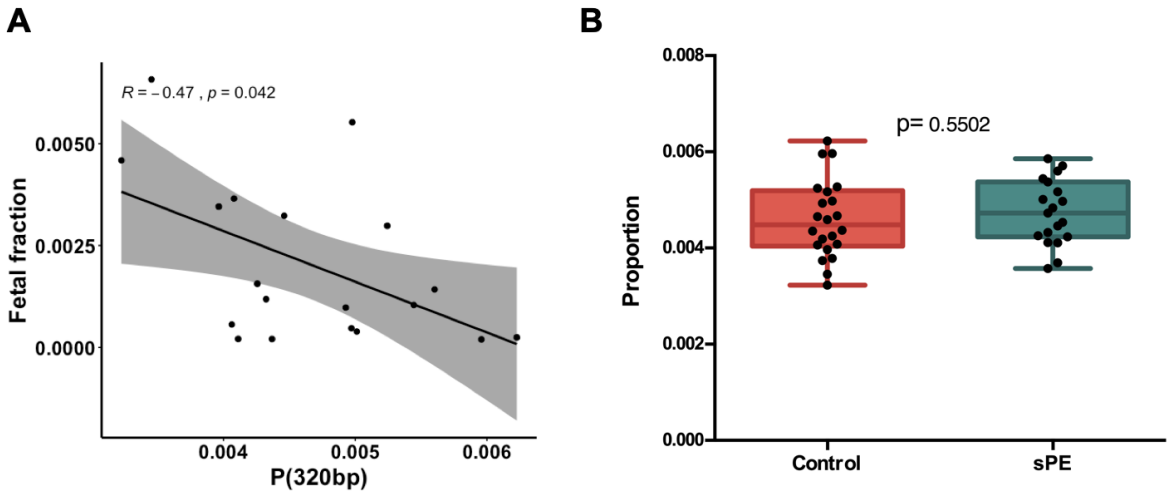

Figure 35. Proportion of 320 bp cell-free DNA fragments analysis. A: Scatter plot shows association between fetal fraction (y-axis) and proportion of 320 bp cfDNA molecules (x-axis) for samples carrying a male fetus (Control n=11, sPE n=8). Linear regression is shown in the plot. R= -0.47, p-value= 0.042. **B**: Box-and-whiskers charts represent proportion of molecules with a size of 320 bp in Control (n=22) and sPE (n=19) samples, p-value= 0.5502. *R= Spearman's correlation coefficient*

VI. DISCUSSION

VI. DISCUSSION

The study of plasma cfDNA has emerged as a valuable source of information that represents a mix of DNA fragments derived from different parts of the body. Its potential in the prognostication, molecular profiling and monitoring of cancer has been demonstrated, leading to the implementation of liquid biopsy in this field (Wan *et al.*, 2017). Another major application of cfDNA is NIPT (Gai and Sun, 2019). However, cfDNA analysis has not yet been applied in other settings. Indeed, assessing epigenetic marks in cfDNA, for example, to trace tissue of origin, is a recent advancement for potentially applying cfDNA analysis beyond cancer and NIPT settings (Snyder *et al.*, 2016; Lehmann-Werman *et al.*, 2016; Moss *et al.*, 2018). Importantly, epigenetic analysis represents an opportunity to detect disease-associated changes that are not based on gene mutations. Thus, analysis of epigenetic biomarkers in cfDNA represents a promising approach for plasma cfDNA-based disease characterization and monitoring that could provide clinical utility in both disease diagnosis and comorbidity studies. Here, we utilized this premise to study whether methylation changes identified in plasma cfDNA can be used for the identification of pathological markers of disease and disease-associated complications.

Several methods are currently available for DNA methylation profiling. Here, MBD-seq was used. This selection reflects that, while bisulfite sequencing represents the gold standard method in the study of DNA methylation, bisulfite treatment produces DNA degradation that complicates its use in low DNA concentration samples. In contrast, methylated DNA enrichment methods represent an optimal approach to cfDNA methylation study since no DNA degradation is produced. In comparison to MeDIP enrichment, which is more sensitive to fragments containing single CpG sites, elution of MBD enrichment protocols allow the study of regions depending on their methylation density (Harris *et al.*, 2010).

Despite the lack of single-base resolution of enrichment methods, these techniques produce a comprehensive profiling of DNA methylation (Taiwo *et al.*, 2012; Harris *et al.*, 2010; Serre *et al.*, 2010) and detect DMRs (Aberg *et al.*, 2012; Nair *et al.*, 2011; Walker *et al.*, 2015). Further, in a comparison between MBD-seq and WGBS, the enrichment approach approximates the sensitivity of WGBS, and specificity was generally better for MBD-seq than WGBS (*DNA*

Methylation Protocols, 2018). Additionally, methylation-enriched DNA sequencing is a cost-effective method for the analysis of cfDNA methylation changes associated to disease. In a clinical landscape, it represents an affordable option that facilitates higher sample throughput than WGBS.

Here, single-read sequencing was used as an affordable and rapid first approach in the discovery of preeclampsia-associated changes. Then, paired-end sequencing was used to obtain a deeper knowledge of DMRs as well as to study cfDNA fragmentation patterns. MBD capture of methylated cfDNA fragments allows not only the identification of DMRs, but also the segregation of DNA methylated reads depending on their fragment size. The combination of these techniques is described for the first time in this thesis, showing the potential of these combined methods for biomarker research.

We applied this approach in an experimental group of pregnant women with and without sPE diagnosis. Preeclampsia is a hypertensive syndrome that affects 2-8% of pregnancies, and the most common hypertensive disorder of pregnancy (ACOG, 2019). The signs observed in women with preeclampsia during the 2nd half of pregnancy are broad and include hypertension, proteinuria, visual disturbances, edema, and liver and kidney failure. Severe forms of preeclampsia can lead to maternal death; indeed, preeclampsia is one of the leading causes of maternal morbidity and mortality worldwide (Ghulmiyyah and Sibai, 2012). We considered that sPE represents a multi-system syndrome model, in which the presence of placenta and the maternal multi-organ involvement could enable the discovery of systemic epigenetic changes associated with the disease, as well as could provide a bird's-eye view of the processes underlying the pathology that could help to improve its understanding.

For this purpose, MBD-captured cfDNA from pregnant women with and without sPE were compared using next generation sequencing. Analysis revealed the presence of DMRs between sPE and control pregnancies, which, to the best of our knowledge, represents the first report of significant differences in cfDNA methylation between these groups.

The fact that distribution of DMRs throughout the genome was significantly different between groups may suggest a potential role of altered cfDNA methylation in the disease. Similar results were obtained in both experimental approaches, where the greatest difference was

observed in promoters. Specifically, the proportion of regions annotated as promoter was significantly higher in regions hypermethylated in sPE (Figures 10 and 18). Furthermore, presence of CpG islands was also significantly overrepresented in hypermethylated regions (Figures 11 and 19), with most of them annotated to gene promoter regions (Figures 12 and 20). If more than 60% of human promoters have CpG islands, most of them are unmethylated during normal development (Baylin and Jones, 2016). Here, the difference in methylated CpG islands detected in our DMRs (hypo- and hyper-methylated) is indicative of the relevance of methylation changes underlying preeclampsia.

Interestingly, hypermethylated regions found in cfDNA had a methylation pattern similar to that observed in cancer cells. The gain of methylated CpG islands in promoters is considered a hallmark of cancer-associated methylation changes (Robertson, 2005). Placentation is often compared to cancer since placenta, like tumors, has to promote angiogenesis and invade tissues while escaping the immune system. Placental cells also share some of their epigenetic fingerprint with cancer cells (Schroeder *et al.*, 2013). Both types of cells are characterized by a widespread hypomethylation and focal hypermethylation at CpG islands. In cancer, hypermethylation at CpG islands in promoter regions of tumor suppressor genes (TSGs) contributes to tumorigenesis by silencing genes involved in cell-cycle regulation or DNA repair (Robertson, 2005). Similarly, hypermethylation of TSGs is observed in healthy placenta as part of its invasive nature; *RASSF1A* hypermethylation is the best-known example (Chiu *et al.*, 2007). Hypermethylation of TSGs, such as negative regulators of Wnt signaling, has been reported in placenta (Novakovic *et al.*, 2008; Yuen *et al.*, 2009). These findings support the involvement of TSG silencing and Wnt signaling in placentation, contributing to pseudo-malignancy of the placenta.

However, if a normal pregnancy is similar to cancer in terms of successful invasion of host tissue, preeclampsia represents the opposite scenario, with impaired angiogenesis (Liu *et al.*, 2015) and impaired evasion of the immune system (Lokki *et al.*, 2018). We therefore hypothesized that the observed differences in CpG island methylation found in cfDNA between hypo- and hypermethylated regions might be present at genes that could explain defective placentation in preeclampsia. Accordingly, we found an enrichment in pathways related to cell cycle, cancer and Wnt signaling for genes in regions hypomethylated in sPE

(Figures 22.A). Previous work also reported a downregulation of Wnt pathway (Wang *et al.*, 2018) and overexpression of TSGs (Heikkilä *et al.*, 2005) in preeclamptic placentas compared to control.

Additionally, functional analysis revealed that genes with hypermethylated CpG island promoters were enriched in several pathways already implicated in preeclampsia pathology, such as MAPK, mTOR, VEGF, calcium, insulin, and adipocytokine signaling pathways; vascular smooth muscle contraction; and cancer (Figures 14 and 22.B). Specifically, VEGF signaling pathway disruption occurs in preeclampsia due to increased placental production of the sFLT1 (Maynard *et al.*, 2003). Inhibition of VEGF signaling by anti-cancer drugs leads to the development of a preeclampsia-like syndrome with hypertension, proteinuria, and renal failure in human (Pandey *et al.*, 2018), which supports a central role of this pathway in the development of preeclampsia symptoms. We also found hypermethylated gene enrichment in the vascular smooth muscle contraction pathway (Figure 22.B). Endothelial cells respond to different stimuli and release molecules that modulate vascular tone, making vascular smooth muscle cells relax or contract. In preeclampsia, systemic endothelial dysfunction leads to an imbalance between relaxing and contracting factors. The increase in vasodilators, coupled with the alteration in responsiveness of vascular smooth muscle cells to these factors, provokes an increase in vasoconstriction responsible for the appearance of hypertension in preeclampsia (Goulopoulou and Davidge, 2015).

In addition to the terms directly related to preeclampsia symptoms, gene-disease association results mirrored some of the long-term consequences of the disease (Figures 16 and 24). Women with preeclampsia have higher risk for developing later-life chronic hypertension, cerebrovascular disease, ischemic heart disease, and stroke, as well as end-stage renal disease (Lykke *et al.*, 2009; Jim and Karumanchi, 2017). Systemic endothelial dysfunction during preeclampsia is proposed to play a role in predisposing to cardiovascular disease. On the other hand, shared risk factors may predispose to preeclampsia as well as to endothelial dysfunction and cardiovascular disease (Powe *et al.*, 2011). Such possibilities could be explained by the nature of DNA methylation. Methylation patterns are transmitted during DNA replication (Probst *et al.*, 2009). This cellular epigenetic memory integrates the information of the local chromatin environment. Perpetuation of DNA methylation patterns is critical for retaining

cellular identity after cell division, which is necessary for maintenance of the phenotype of differentiated cells. Changes incurred during a disease could be linked to future sequelae by perpetuation of epigenetic modifications. Future work is needed to identify potential targets that may help to understand later-life pathologies associated to past sPE. Likewise, hypermethylated genes are significantly enriched in type 2 diabetes; women who have had preeclampsia are at increased risk for developing this disease compared to normotensive pregnancies (Lykke *et al.*, 2009). An additional link between type 2 diabetes and preeclampsia is the utility of Metformin, a drug used for diabetes treatment that can also prevent hypertensive disorders of pregnancy, including preeclampsia (Kalafat *et al.*, 2018; Alqudah *et al.*, 2018).

Interestingly, changes were also associated with differential TFBS methylation. Specifically, TFBS over-represented among hypomethylated regions corresponded to TFs enriched in processes implicated in positive regulation of transcription, gene expression, and metabolic and biosynthetic processes (Figure 15.A and 23.A) and TFs corresponding to TFBS hypermethylated in sPE were significantly associated with negative regulation processes (Figure 15.B and 23.B). Although differences are clear in our experimental set-up, functional analysis of this phenomenon is not straight-forward. Until recently, DNA methylation was thought to prevent TF binding as in the case of CTCF (Bell and Felsenfeld, 2000). However, some works have demonstrated that this cannot be considered a rule. An *in vitro* study from Yin *et al.* found that DNA methylation prevented TF binding in some cases but it enhanced TF binding in others. Moreover, they also found some TFs that were not affected by DNA methylation (Yin *et al.*, 2017). Similarly, a study comparing ChIP-seq and WGBS data from different cell types showed that DNA methylation effects on TFs binding are variable among different TFs (Xuan Lin *et al.*, 2019). Therefore, although these transcriptionally important regions showed differences in methylation, we cannot infer the effects these changes in methylation of TFBS have on gene expression.

In conclusion, functional analysis results correlated to changes previously observed in preeclampsia. Although we cannot infer gene expression changes from these data, we conclude that cfDNA methylation differences between sPE and control pregnancies reflect changes/alterations in pathways previously related to the disease as well as its main

symptoms. Our results highlight, for the first time, the potential of methylome cfDNA analysis in the study of sPE.

We also explored cfDNA fragmentation patterns. cfDNA fragmentation is not random; in fact, molecule size correlates with nucleosome footprint and overall distribution reflects apoptotic fragmentation. cfDNA originating from different tissues present differences in molecules size (Zheng *et al.*, 2012), as well as differential cutting sites (Sun *et al.*, 2018a). Nucleosome occupancy is highly related to cell identity, and differential chromatin accessibility is thought to be the main contributing factor for the observed differences (Snyder *et al.*, 2016; Sun *et al.*, 2018a). Thus, the type of cell, nucleosome positioning, and mechanisms of cfDNA release can mark molecules with specific fragmentation signatures that potentially provide information about cfDNA origin (Snyder *et al.*, 2016), gene expression (Ulz *et al.*, 2016; Ivanov *et al.*, 2015), or the existence of pathology (Jiang *et al.*, 2015). Recent studies reveal the utility of cfDNA fragmentation patterns as an added source of information in NIPT (Yu *et al.*, 2014; Hellwig *et al.*, 2018; Mouliere *et al.*, 2018). Based on these data, we aimed to study fragment distribution in sPE and control pregnancies. The multi-organ damage characteristic of the maternal syndrome of sPE, along with placental involvement in its pathology, should be mirrored by cfDNA. We hypothesized that cfDNA fragmentation differences between sPE and control pregnancies could reflect differential tissue contribution to methylated cfDNA.

An enrichment in short fragments was observed in preeclamptic pregnancies. This enrichment was detected in the fragments around the 1-nucleosome peak, but also in the 2-nucleosome peak (Figure 27). We hypothesized that the observed increase in short fragments must be due either to an increase in fetal cfDNA (Lo *et al.*, 2010) or to an increase in non-hematopoietically derived cfDNA from maternal organs and tissues (Zheng *et al.*, 2012). To answer this question, fetal contribution was estimated; a lower fetal fraction was observed in sPE pregnancies, but differences did not reach statistical significance (Figures 17 and 26). To further investigate if the observed differences in fragment size between groups were due to differences in fetal contribution to cfDNA, association with each variable was assessed. Surprisingly, we found a negative correlation between “short” fragments (270-320 bp) and fetal fraction, and positive correlation between “long” fragments (175-225 bp) and fetal fraction (Figure 33). These data suggest that the observed increase in short fragments in women with preeclampsia is not due

to an increase in fetal cfDNA. Supporting that, a significant association was found between the proportion of 320-bp fragments and fetal fraction, but no differences were observed between sPE and control pregnancies (Figure 35). This observation underscores that the observed differences in fragment size between groups are not due to differential fetal contributions.

In view of these data, we hypothesized that the increase in methylated short fragments should come from non-hematopoietic tissues that might originate from maternal organs in response to preeclampsia pathology. This hypothesis aligns with published studies examining fetal contribution in preeclampsia, in which the increase in fetal cfDNA (Lo *et al.*, 1999; Martin *et al.*, 2014; Contro *et al.*, 2017) is accompanied by a larger increase in maternal cfDNA, causing a reduction in fetal fraction (Rolnik *et al.*, 2015; Rolnik *et al.*, 2018; Bender *et al.*, 2019; Gerson *et al.*, 2019). Notably, size differences between cfDNA molecules were previously described using total cfDNA; in our study, only methylated cfDNA was analyzed. In the relationship between methylation and cfDNA fragment size, DNA molecules with high methylation densities were observed to be longer than the ones with lower densities (Lun *et al.*, 2013). Lun *et al.* observed the same pattern in maternal and fetal cfDNA molecules in pregnant women, but also in cfDNA from non-pregnant women. This observation suggests that the hypomethylated state of fetal DNA is not the only factor that accounts for its relative shortness compared with maternal DNA, and other factors that can be associated to fetal origin may also be relevant, at least in a pathological situation. Based on that, we conclude that the analysis of methylated DNA only (compared to total cfDNA) would enable successful study of fragment distributions. Although further investigation using tissue-specific methylation marks is needed to confirm our findings, these data suggest that observed differences in methylated cfDNA size may reflect differential tissue contribution in sPE and control pregnancies.

Here, for the first time, methylated cfDNA fragment size analysis has revealed differences between sPE and control pregnancies. Current protocols for NIPT use paired-end sequencing, which allows fragment size analysis without any additional steps during library preparation or sequencing. While further research is needed to confirm our findings in early pregnancy stages, our results suggest that fragment size analysis would represent a valuable source of information to estimate preeclampsia risk during NIPT. In fact, low fetal fraction has already

been used to estimate risks of preeclampsia and other complications during in NIPT (Rolnik *et al.*, 2018; Gerson *et al.*, 2019).

Based on our findings, which demonstrate the potential of cfDNA methylome analysis for the study of mechanisms underlying sPE, we further explored the presence of gene-specific methylation changes. It is important to keep in mind that the observed DMRs can be due to either disease-specific DNA methylation changes in particular tissues, or to changes in the proportions of cfDNA from different tissues (Feng *et al.*, 2018). Although we cannot discard that some of the DMRs can result from differences in tissue origin, our results showed potentially relevant changes in genes previously related to preeclampsia (Annex 3).

Among hypermethylated regions we found DMRs at genes involved in angiogenesis and vascular tone regulation, such as endoglin (*ENG*). The role of its soluble form (sENG) in preeclampsia pathogenesis is already known (Venkatesha *et al.*, 2006) (Figure 3), and sENG protein levels have been related to preeclampsia (ACOG, 2013). Apart from its role in angiogenesis and vascular tone regulation, ENG negatively regulates trophoblast differentiation towards the invasive phenotype (Caniggia *et al.*, 1997; Mano *et al.*, 2011). Its expression is increased in preeclamptic placentas (Venkatesha *et al.*, 2006; Sitras *et al.*, 2009), in maternal blood cells (Purwosunu *et al.*, 2009; Farina *et al.*, 2010), and in chorionic villous samples of women who later developed preeclampsia (Farina *et al.*, 2008). Observed differences in *ENG* cfDNA methylation could reflect differential gene regulation in sPE and control pregnancies.

In addition, among regions hypermethylated in sPE, two DMRs were present in the *DNMTA3* gene. This DNA methyltransferase, which catalyzes de-novo methylation, has been previously related to preeclampsia. Specifically, *DNMTA3* expression was lower in hypoxic trophoblasts and preeclamptic placentas when compared with normoxic trophoblasts and healthy placentas, respectively (Ma *et al.*, 2018). Moreover, its hypermethylation has been suggested to be related to preeclampsia due to immunological rejection (Leavey *et al.*, 2018).

We also found one DMR in the *MMP23B* gene, a matrix metalloprotease (MMP). MMPs are important regulators of vascular and uterine remodeling and they play a major role in

trophoblast invasion into the uterine wall. *MMP23B* expression varies across the pregnancy in decidua, where it is expressed in trophoblasts, uterine NK cells, decidual stroma cells and decidual glands (Anacker *et al.*, 2011). For that reason, Anacker *et.* al suggested a possible role of some MMPs, including *MMP23B*, in the cross-talk between mother and embryo during early pregnancy. A role for MMPs in preeclampsia has been widely studied. Preeclampsia is associated with a decreased expression of *MMP-2* and *-9*, as well as an increase in their circulating levels (Espino y Sosa *et al.*, 2017). Moreover, DNA methylation differences at some genes encoding MMPs have been observed between preeclamptic and healthy placentas (Wang *et al.*, 2010; Rahat *et al.*, 2016a) and between omental arteries from normal pregnant women and those with preeclampsia (Mousa *et al.*, 2012b). These data suggest an epigenetic regulation that may account for the observed differences in MMP expression and circulating levels.

Moreover, one DMR was also present at the *VEGFB* gene, another member of the VEGF family that binds to FMS-like tyrosine kinase (FLT1) or to its soluble form (sFLT1). VEGFB stimulates vessel growth primarily in pathological conditions (Li *et al.*, 2008). Administration of exogenous growth factors that bind FLT-1 and sFLT-1 have been proposed as strategies to treat maternal syndrome of preeclampsia; VEGF has demonstrated therapeutic efficacy in this setting (Eddy *et al.*, 2018). However, the potent activity of VEGF is associated with adverse effects. Unlike VEGF, VEGFB does not bind to FLK-1 receptor tyrosine kinase (FLK1), which reduces the risk of adverse effects because of aberrant angiogenesis caused by VEGF (Logue *et al.*, 2017). For this reason, VEGFB has been proposed as a safer therapeutic agent for preeclampsia treatment, which would not induce unwanted angiogenesis and increased vascular permeability like that observed with administration of the isoform VEGF-A (Waller *et al.*, 2019).

Additionally, we found three hypermethylated regions in the *eNOS* gene (Figure 25); eNOS plays a known role in the regulation of vascular smooth tone. Endothelial cell production of nitric oxide (NO) leads to vascular smooth muscle relaxation. In preeclampsia, endothelial dysfunction results in a decrease in NO production (Wang *et al.*, 2009) and subsequent vasoconstriction. Alterations of this gene function may have an important role in preeclampsia symptoms since blocking the production of NO induces the majority of symptoms of PE in

pregnant rats (Molnár *et al.*, 1994), and genetic variations in the *eNOS* gene contribute to an increased risk for preeclampsia in humans (Zeng *et al.*, 2016). Recent data based on results from preclinical studies and clinical trials shows that enhancing NO-mediated effects using sildenafil improves fetal growth and maternal blood pressure regulation in preeclampsia (Paauw *et al.*, 2017).

Expression of *eNOS* is mostly restricted to endothelial cells (Förstermann and Sessa, 2012), and DNA methylation has a proven role in cell-specific constitutive expression of *eNOS* in the vascular endothelium (Chan *et al.*, 2004; Shirodkar *et al.*, 2013). Methylation has been also been related to the differential expression of *eNOS* between placenta-derived venous and arterial endothelial cells from uncomplicated pregnancies (Joo *et al.*, 2013). In that study, two regions showed significant differences in *eNOS* DNA methylation. One region, hypermethylated in endothelial cells from placental veins, was located in the *eNOS* major promoter; the second region, hypermethylated in endothelial cells from placental arteries, was located in a downstream exon-associated region overlapping a CpG island (Joo *et al.*, 2013). Joo *et al.* attributed differential *eNOS* expression to gene promoter hypermethylation since it was more expressed in cells from placental arteries. However, the possible regulatory function of the second region was not determined.

Some works in cancer have demonstrated that CpG island methylation within a gene body results in gene overexpression (Sun *et al.*, 2018c; Arechederra *et al.*, 2018); similar observations hold in non-cancerous scenarios such as changes from early- to late-pregnancy placentas (Lim *et al.*, 2017). In this case, CpG island hypermethylation at the downstream region might be implicated in the differential expression observed between endothelial cells from veins and arteries in Joo *et al.*'s work (Joo *et al.*, 2013). Further investigation and additional data will be required to elucidate the role of CpG island gene body hypermethylation in *eNOS* gene expression regulation. Surprisingly, we found two hypermethylated regions, one overlapping the same CpG island (Figure 25) and one located at the corresponding CpG shore. Based on the observed differences in cfDNA methylation and in the potential implication of *eNOS* CpG island methylation in gene expression regulation, we suggest that it represents an interesting marker to study further pathological features in preeclamptic pregnancies.

To our knowledge, the present study represents the first report of significant differences in cfDNA methylation between preeclamptic and healthy pregnancies. Previous works have studied cfDNA and preeclampsia, but focus has been put on cfDNA concentration differences between healthy pregnancies and patients with preeclampsia (Lo *et al.*, 1999; Martin *et al.*, 2014; Contro *et al.*, 2017; Rolnik *et al.*, 2018; Rolnik *et al.*, 2015; Bender *et al.*, 2019; Gerson *et al.*, 2019). In this work, we went a step further and explored the hypothesis that cfDNA in sPE patients may carry substantial biological information with relevance in the sPE pathological signature.

We found several DMRs with a potential functional effect on disease pathology. In our context it is important to consider that it is not possible rule out whether some DMRs arise from differences in tissue origin. Nevertheless, the identification by gene-disease ontology of several pathological features in sPE supports our hypothesis. Additionally, for the first time, methylated cfDNA fragment size analysis has revealed differences between sPE and control pregnancies. Our data underscore the utility of cfDNA fragment analysis in sPE studies, which could shed light on differential tissue contributions.

cfDNA is released mainly from dying cells during the process of apoptosis; the fact that it originates from different tissues gives an opportunity to study whole-body status. Moreover, epigenetic information represents a valuable source of information in scenarios not involving DNA mutations. Our disease model, preeclampsia, presents a complex pathophysiology. It affects not only fetal, but also maternal health: hypertension, impaired liver function, and systemic inflammatory responses are observed in women with preeclampsia. For this reason, we considered that it represents the perfect scenario to assess the power of low-pass cfDNA MBD-seq in the search for specific epigenetic changes that could provide information about health status. Here, our hypothesis that cfDNA methylation would mirror preeclampsia systemic responses has been supported by empirical evidence.

Our results highlight the potential utility of this simple method for other disease settings. Additional research is warranted to establish cfDNA methylomes in both controls and patients to reveal patterns linked to individual patient pathology, to ultimately improve diagnosis,

treatment and prevention, with the potential to be used not only in preeclampsia, but also in other clinical scenarios.

VII. CONCLUSIONS

VII. CONCLUSIONS

- MBD-sequencing of cfDNA revealed the presence of differentially methylated regions between control pregnancies and women with severe preeclampsia.

- The enrichment in promoters and CpG islands among hypermethylated regions supports the role of DNA methylation during mechanisms underlying severe preeclampsia.

- cfDNA methylation differences between severe preeclampsia and control pregnancies reflect alterations in pathways previously related to the disease, such as pathways in cancer, Wnt signaling, VEGF signaling, and vascular smooth muscle contraction.

- Gene-disease association analysis of hypermethylated genes showed an enrichment in genes related to preeclampsia symptoms and long-term consequences, suggesting a potential methylation involvement in the mechanisms that predispose women with preeclampsia to cardiovascular and metabolic diseases later in life.

- Fragment size analysis revealed significant differences between control and severe preeclamptic pregnancies, with an enrichment in short fragments in sPE negatively correlated with fetal fraction. These data point to maternal increased contribution to cfDNA in response to the pathological condition.

- MBD-captured cfDNA analysis allows the identification of methylation differences in a clinical diagnostic work flow.

- The study of epigenetic features in cell-free DNA represents an exploitable source of information in diseases with multi-organ involvement such as severe preeclampsia. Specifically, the use of epigenetic differences widens liquid biopsy application for biomarker detection in diseases without high mutation rates.

VIII. BIBLIOGRAPHY

VII. BIBLIOGRAPHY

Aberg, K. A., McClay, J. L., Nerella, S., Xie, L. Y., Clark, S. L., Hudson, A. D., Bukszár, J., Adkins, D., Hultman, C. M., Sullivan, P. F., Magnusson, P. K., van den Oord, E. J. and Consortium, S. S. (2012) 'MBD-seq as a cost-effective approach for methylome-wide association studies: demonstration in 1500 case--control samples', *Epigenomics,* 4(6), pp. 605-21.

ACOG (2013) 'Hypertension in pregnancy', *Obstet Gynecol,* 122(5), pp. 1122-31.

ACOG (2019) 'ACOG Practice Bulletin No. 202', *Obstet Gynecol,* 133(1), pp. e1-e25.

Acosta-Sison, H. (1956) 'The relationship of hydatidiform mole to pre-eclampsia and eclampsia; a study of 85 cases', *Am J Obstet Gynecol,* 71(6), pp. 1279-82.

Alqudah, A., McKinley, M. C., McNally, R., Graham, U., Watson, C. J., Lyons, T. J. and McClements, L. (2018) 'Risk of pre-eclampsia in women taking metformin: a systematic review and meta-analysis', *Diabet Med,* 35(2), pp. 160-172.

Anacker, J., Segerer, S. E., Hagemann, C., Feix, S., Kapp, M., Bausch, R. and Kämmerer, U. (2011) 'Human decidua and invasive trophoblasts are rich sources of nearly all human matrix metalloproteinases', *Mol Hum Reprod,* 17(10), pp. 637-52.

Anderson, C. M., Ralph, J. L., Johnson, L., Scheett, A., Wright, M. L., Taylor, J. Y., Ohm, J. E. and Uthus, E. (2015) 'First trimester vitamin D status and placental epigenomics in preeclampsia among Northern Plains primiparas', *Life Sci,* 129, pp. 10-5.

Anderson, C. M., Ralph, J. L., Wright, M. L., Linggi, B. and Ohm, J. E. (2014) 'DNA methylation as a biomarker for preeclampsia', *Biol Res Nurs,* 16(4), pp. 409-20.

Anton, L., Brown, A. G., Bartolomei, M. S. and Elovitz, M. A. (2014) 'Differential methylation of genes associated with cell adhesion in preeclamptic placentas', *PLoS One,* 9(6), pp. e100148.

Arechederra, M., Daian, F., Yim, A., Bazai, S. K., Richelme, S., Dono, R., Saurin, A. J., Habermann, B. H. and Maina, F. (2018) 'Hypermethylation of gene body CpG islands predicts high dosage of functional oncogenes in liver cancer', *Nat Commun,* 9(1), pp. 3164.

Aucamp, J., Bronkhorst, A. J., Badenhorst, C. P. S. and Pretorius, P. J. (2018) 'The diverse origins of circulating cell-free DNA in the human body: a critical re-evaluation of the literature', *Biol Rev Camb Philos Soc,* 93(3), pp. 1649-1683.

Barlow, D. P. and Bartolomei, M. S. (2014) 'Genomic imprinting in mammals', *Cold Spring Harb Perspect Biol,* 6(2).

Barros-Silva, D., Marques, C. J., Henrique, R. and Jerónimo, C. (2018) 'Profiling DNA Methylation Based on Next-Generation Sequencing Approaches: New Insights and Clinical Applications', *Genes (Basel),* 9(9).

Bartolomei, M. S. and Ferguson-Smith, A. C. (2011) 'Mammalian genomic imprinting', *Cold Spring Harb Perspect Biol,* 3(7).

Baylin, S. B. and Jones, P. A. (2016) 'Epigenetic Determinants of Cancer', *Cold Spring Harb Perspect Biol,* 8(9).

Becker, K. G., Barnes, K. C., Bright, T. J. and Wang, S. A. (2004) 'The genetic association database', *Nat Genet: Vol. 5*. United States, pp. 431-2.

Bell, A. C. and Felsenfeld, G. (2000) 'Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene', *Nature,* 405(6785), pp. 482-5.

Bender, W. R., Koelper, N. C., Sammel, M. D. and Dugoff, L. (2019) 'Association of Fetal Fraction of Cell-Free DNA and Hypertensive Disorders of Pregnancy', *Am J Perinatol,* 36(3), pp. 311-316.

Bergman, Y. and Cedar, H. (2013) 'DNA methylation dynamics in health and disease', *Nat Struct Mol Biol,* 20(3), pp. 274-81.

Bhasin, J. M., Hu, B. and Ting, A. H. (2016) 'MethylAction: detecting differentially methylated regions that distinguish biological subtypes', *Nucleic Acids Res,* 44(1), pp. 106-16.

Bhasin, J. M. and Ting, A. H. (2016) 'Goldmine integrates information placing genomic ranges into meaningful biological contexts', *Nucleic Acids Res,* 44(12), pp. 5550-6.

Bianchi, D. W., Chudova, D., Sehnert, A. J., Bhatt, S., Murray, K., Prosen, T. L., Garber, J. E., Wilkins-Haug, L., Vora, N. L., Warsof, S., Goldberg, J., Ziainia, T. and Halks-Miller, M. (2015) 'Noninvasive Prenatal Testing and Incidental Detection of Occult Maternal Malignancies', *JAMA,* 314(2), pp. 162-9.

Blair, J. D., Yuen, R. K., Lim, B. K., McFadden, D. E., von Dadelszen, P. and Robinson, W. P. (2013) 'Widespread DNA hypomethylation at gene enhancer regions in placentas associated with early-onset pre-eclampsia', *Mol Hum Reprod,* 19(10), pp. 697-708.

Burton, G. J., Watson, A. L., Hempstock, J., Skepper, J. N. and Jauniaux, E. (2002) 'Uterine glands provide histiotrophic nutrition for the human fetus during the first trimester of pregnancy', *J Clin Endocrinol Metab,* 87(6), pp. 2954-9.

Caniggia, I., Grisaru-Gravnosky, S., Kuliszewsky, M., Post, M. and Lye, S. J. (1999) 'Inhibition of TGF-beta 3 restores the invasive capability of extravillous trophoblasts in preeclamptic pregnancies', *J Clin Invest,* 103(12), pp. 1641-50.

Caniggia, I., Mostachfi, H., Winter, J., Gassmann, M., Lye, S. J., Kuliszewski, M. and Post, M. (2000) 'Hypoxia-inducible factor-1 mediates the biological effects of oxygen on human trophoblast differentiation through TGFbeta(3)', *J Clin Invest,* 105(5), pp. 577-87.

Caniggia, I., Taylor, C. V., Ritchie, J. W., Lye, S. J. and Letarte, M. (1997) 'Endoglin regulates trophoblast differentiation along the invasive pathway in human placental villous explants', *Endocrinology,* 138(11), pp. 4977-88.

Cedar, H. and Bergman, Y. (2012) 'Programming of DNA methylation patterns', *Annu Rev Biochem,* 81, pp. 97-117.

Chaiworapongsa, T., Chaemsaithong, P., Yeo, L. and Romero, R. (2014) 'Pre-eclampsia part 1: current understanding of its pathophysiology', *Nat Rev Nephrol,* 10(8), pp. 466-80.

Chan, K. C., Jiang, P., Sun, K., Cheng, Y. K., Tong, Y. K., Cheng, S. H., Wong, A. I., Hudecova, I., Leung, T. Y., Chiu, R. W. and Lo, Y. M. (2016) 'Second generation noninvasive fetal genome analysis reveals de novo mutations, single-base parental inheritance, and preferred DNA ends', *Proc Natl Acad Sci U S A,* 113(50), pp. E8159-E8168.

Chan, Y., Fish, J. E., D'Abreo, C., Lin, S., Robb, G. B., Teichert, A. M., Karantzoulis-Fegaras, F., Keightley, A., Steer, B. M. and Marsden, P. A. (2004) 'The cell-specific expression of endothelial nitric-oxide synthase: a role for DNA methylation', *J Biol Chem,* 279(33), pp. 35087-100.

Chelbi, S. T., Mondon, F., Jammes, H., Buffat, C., Mignot, T. M., Tost, J., Busato, F., Gut, I., Rebourcet, R., Laissue, P., Tsatsaris, V., Goffinet, F., Rigourd, V., Carbonne, B., Ferré, F. and Vaiman, D. (2007) 'Expressional and epigenetic alterations of placental serine protease inhibitors: SERPINA3 is a potential marker of preeclampsia', *Hypertension,* 49(1), pp. 76-83.

Chen, Z., Li, S., Subramaniam, S., Shyy, J. Y. and Chien, S. (2017) 'Epigenetic Regulation: A New Frontier for Biomedical Engineers', *Annu Rev Biomed Eng,* 19, pp. 195-219.

Chiu, R. W., Chim, S. S., Wong, I. H., Wong, C. S., Lee, W. S., To, K. F., Tong, J. H., Yuen, R. K., Shum, A. S., Chan, J. K., Chan, L. Y., Yuen, J. W., Tong, Y. K., Weier, J. F., Ferlatte, C., Leung, T. N., Lau, T. K., Lo, K. W. and Lo, Y. M. (2007) 'Hypermethylation of RASSF1A in human and rhesus placentas', *Am J Pathol,* 170(3), pp. 941-50.

Chu, T., Bunce, K., Shaw, P., Shridhar, V., Althouse, A., Hubel, C. and Peters, D. (2014) 'Comprehensive analysis of preeclampsia-associated DNA methylation in the placenta', *PLoS One,* 9(9), pp. e107318.

Contro, E., Bernabini, D. and Farina, A. (2017) 'Cell-Free Fetal DNA for the Prediction of Pre-Eclampsia at the First and Second Trimesters: A Systematic Review and Meta-Analysis', *Mol Diagn Ther,* 21(2), pp. 125-135.

Corcoran, R. B. and Chabner, B. A. (2018) 'Application of Cell-free DNA Analysis to Cancer Treatment', *N Engl J Med,* 379(18), pp. 1754-1765.

Cunningham, F. G., Leveno, K. J., Bloom, S. L., Spong, C. Y., Dashe, J. S. and Hoffman, B. L. (2014) *Williams Obstetrics 24th Edition.* 24th edn.: McGraw-Hill Education.

Deaton, A. M. and Bird, A. (2011) 'CpG islands and the regulation of transcription', *Genes Dev,* 25(10), pp. 1010-22.

DNA Methylation Protocols. (2018) 3rd edn.

Dor, Y. and Cedar, H. (2018) 'Principles of DNA methylation and their implications for biology and medicine', *Lancet,* 392(10149), pp. 777-786.

Doridot, L., Houry, D., Gaillard, H., Chelbi, S. T., Barbaux, S. and Vaiman, D. (2014) 'miR-34a expression, epigenetic regulation, and function in human placental diseases', *Epigenetics,* 9(1), pp. 142-51.

Duley, L. (2009) 'The global impact of pre-eclampsia and eclampsia', *Semin Perinatol,* 33(3), pp. 130-7.

Eddy, A. C., Bidwell, G. L. and George, E. M. (2018) 'Pro-angiogenic therapeutics for preeclampsia', *Biol Sex Differ,* 9(1), pp. 36.

Espino y Sosa, S., Flores-Pliego, A., Espejel-Nunez, A., Medina-Bastidas, D., Vadillo-Ortega, F., Zaga-Clavellina, V. and Estrada-Gutierrez, G. (2017) 'New Insights into the Role of Matrix Metalloproteinases in Preeclampsia', *Int J Mol Sci,* 18(7).

Farina, A., Sekizawa, A., De Sanctis, P., Purwosunu, Y., Okai, T., Cha, D. H., Kang, J. H., Vicenzi, C., Tempesta, A., Wibowo, N., Valvassori, L. and Rizzo, N. (2008) 'Gene expression in chorionic villous samples at 11 weeks' gestation from women destined to develop preeclampsia', *Prenat Diagn,* 28(10), pp. 956-61.

Farina, A., Zucchini, C., Sekizawa, A., Purwosunu, Y., de Sanctis, P., Santarsiero, G., Rizzo, N., Morano, D. and Okai, T. (2010) 'Performance of messenger RNAs circulating in maternal blood in the prediction of preeclampsia at 10-14 weeks', *Am J Obstet Gynecol,* 203(6), pp. 575.e1-7.

Feng, H., Jin, P. and Wu, H. (2018) 'Disease prediction by cell-free DNA methylation', *Brief Bioinform*.

Fenouil, R., Cauchy, P., Koch, F., Descostes, N., Cabeza, J. Z., Innocenti, C., Ferrier, P., Spicuglia, S., Gut, M., Gut, I. and Andrau, J. C. (2012) 'CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters', *Genome Res,* 22(12), pp. 2399-408.

Fernandez, A. F., Assenov, Y., Martin-Subero, J. I., Balint, B., Siebert, R., Taniguchi, H., Yamamoto, H., Hidalgo, M., Tan, A. C., Galm, O., Ferrer, I., Sanchez-Cespedes, M., Villanueva, A., Carmona, J., Sanchez-Mut, J. V., Berdasco, M., Moreno, V., Capella, G., Monk, D., Ballestar, E., Ropero, S., Martinez, R., Sanchez-Carbayo, M., Prosper, F., Agirre, X., Fraga, M. F., Graña, O., Perez-Jurado, L., Mora, J., Puig, S., Prat, J., Badimon, L., Puca, A. A., Meltzer, S. J., Lengauer, T., Bridgewater, J., Bock, C. and Esteller, M. (2012) 'A DNA methylation fingerprint of 1628 human samples', *Genome Res,* 22(2), pp. 407-19.

Ferreira, H. J. and Esteller, M. (2018) 'CpG Islands in Cancer: Heads, Tails, and Sides', *Methods Mol Biol,* 1766, pp. 49-80.

Founds, S. A., Shi, H., Conley, Y. P., Jeyabalan, A., Roberts, J. M. and Lyons-Weiler, J. (2012) 'Variations in discovery-based preeclampsia candidate genes', *Clin Transl Sci,* 5(4), pp. 333-9.

Förstermann, U. and Sessa, W. C. (2012) 'Nitric oxide synthases: regulation and function', *Eur Heart J,* 33(7), pp. 829-37, 837a-837d.

Gai, W. and Sun, K. (2019) 'Epigenetic Biomarkers in Cell-Free DNA and Applications in Liquid Biopsy', *Genes (Basel),* 10(1).

Gao, W. L., Li, D., Xiao, Z. X., Liao, Q. P., Yang, H. X., Li, Y. X., Ji, L. and Wang, Y. L. (2011) 'Detection of global DNA methylation and paternally imprinted H19 gene methylation in preeclamptic placentas', *Hypertens Res,* 34(5), pp. 655-61.

Genereux, D. P., Johnson, W. C., Burden, A. F., Stöger, R. and Laird, C. D. (2008) 'Errors in the bisulfite conversion of DNA: modulating inappropriate- and failed-conversion frequencies', *Nucleic Acids Res,* 36(22), pp. e150.

Gerson, K. D., Truong, S., Haviland, M. J., O'Brien, B. M., Hacker, M. R. and Spiel, M. H. (2019) 'Low fetal fraction of cell-free DNA predicts placental dysfunction and hypertensive disease in pregnancy', *Pregnancy Hypertens,* 16, pp. 148-153.

Ghulmiyyah, L. and Sibai, B. (2012) 'Maternal mortality from preeclampsia/eclampsia', *Semin Perinatol,* 36(1), pp. 56-9.

Goulopoulou, S. and Davidge, S. T. (2015) 'Molecular mechanisms of maternal vascular dysfunction in preeclampsia', *Trends Mol Med,* 21(2), pp. 88-97.

Greally, J. M. (2018) 'A user's guide to the ambiguous word 'epigenetics'', *Nat Rev Mol Cell Biol,* 19(4), pp. 207-208.

Han, H., Cortez, C. C., Yang, X., Nichols, P. W., Jones, P. A. and Liang, G. (2011) 'DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter', *Hum Mol Genet,* 20(22), pp. 4299-310.

Harris, R. A., Wang, T., Coarfa, C., Nagarajan, R. P., Hong, C., Downey, S. L., Johnson, B. E., Fouse, S. D., Delaney, A., Zhao, Y., Olshen, A., Ballinger, T., Zhou, X., Forsberg, K. J., Gu, J., Echipare, L., O'Geen, H., Lister, R., Pelizzola, M., Xi, Y., Epstein, C. B., Bernstein, B. E., Hawkins, R. D., Ren, B., Chung, W. Y., Gu, H., Bock, C., Gnirke, A., Zhang, M. Q., Haussler, D., Ecker, J. R., Li, W., Farnham, P. J., Waterland, R. A., Meissner, A., Marra, M. A., Hirst, M., Milosavljevic, A. and Costello, J. F. (2010) 'Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications', *Nat Biotechnol,* 28(10), pp. 1097-105.

Hartung, T., Zhang, L., Kanwar, R., Khrebtukova, I., Reinhardt, M., Wang, C., Therneau, T. M., Banck, M. S., Schroth, G. P. and Beutler, A. S. (2012) 'Diametrically opposite methylome-transcriptome relationships in high- and low-CpG promoter genes in postmitotic neural rat tissue', *Epigenetics,* 7(5), pp. 421-8.

He, J., Zhang, A., Fang, M., Fang, R., Ge, J., Jiang, Y., Zhang, H., Han, C., Ye, X., Yu, D., Huang, H., Liu, Y. and Dong, M. (2013) 'Methylation levels at IGF2 and GNAS DMRs in infants born to preeclamptic pregnancies', *BMC Genomics,* 14, pp. 472.

Heikkilä, A., Tuomisto, T., Häkkinen, S. K., Keski-Nisula, L., Heinonen, S. and Ylä-Herttuala, S. (2005) 'Tumor suppressor and growth regulatory genes are overexpressed in severe early-onset preeclampsia--an array study on case-specific human preeclamptic placental tissue', *Acta Obstet Gynecol Scand,* 84(7), pp. 679-89.

Hellman, A. and Chess, A. (2007) 'Gene body-specific methylation on the active X chromosome', *Science,* 315(5815), pp. 1141-3.

Hellwig, S., Nix, D. A., Gligorich, K. M., O'Shea, J. M., Thomas, A., Fuertes, C. L., Bhetariya, P. J., Marth, G. T., Bronner, M. P. and Underhill, H. R. (2018) 'Automated size selection for short cell-free DNA fragments enriches for circulating tumor DNA and improves error correction during next generation sequencing', *PLoS One,* 13(7), pp. e0197333.

Herzog, E. M., Eggink, A. J., Willemsen, S. P., Slieker, R. C., Wijnands, K. P. J., Felix, J. F., Chen, J., Stubbs, A., van der Spek, P. J., van Meurs, J. B. and Steegers-Theunissen, R. P. M. (2017) 'Early- and late-onset preeclampsia and the tissue-specific epigenome of the placenta and newborn', *Placenta,* 58, pp. 122-132.

Hiby, S. E., Apps, R., Sharkey, A. M., Farrell, L. E., Gardner, L., Mulder, A., Claas, F. H., Walker, J. J., Redman, C. W., Redman, C. C., Morgan, L., Tower, C., Regan, L., Moore, G. E.,

Carrington, M. and Moffett, A. (2010) 'Maternal activating KIRs protect against human reproductive failure mediated by fetal HLA-C2', *J Clin Invest,* 120(11), pp. 4102-10.

Hiby, S. E., Walker, J. J., O'shaughnessy, K. M., Redman, C. W., Carrington, M., Trowsdale, J. and Moffett, A. (2004) 'Combinations of maternal KIR and fetal HLA-C genes influence the risk of preeclampsia and reproductive success', *J Exp Med,* 200(8), pp. 957-65.

Hu, W., Weng, X., Dong, M., Liu, Y., Li, W. and Huang, H. (2014) 'Alteration in methylation level at 11β-hydroxysteroid dehydrogenase type 2 gene promoter in infants born to preeclamptic women', *BMC Genet,* 15, pp. 96.

Ivanov, M., Baranova, A., Butler, T., Spellman, P. and Mileyko, V. (2015) 'Non-random fragmentation patterns in circulating cell-free DNA reflect epigenetic regulation', *BMC Genomics,* 16 Suppl 13, pp. S1.

Jauniaux, E., Watson, A. L., Hempstock, J., Bao, Y. P., Skepper, J. N. and Burton, G. J. (2000) 'Onset of maternal arterial blood flow and placental oxidative stress. A possible factor in human early pregnancy failure', *Am J Pathol,* 157(6), pp. 2111-22.

Jeziorska, D. M., Murray, R. J. S., De Gobbi, M., Gaentzsch, R., Garrick, D., Ayyub, H., Chen, T., Li, E., Telenius, J., Lynch, M., Graham, B., Smith, A. J. H., Lund, J. N., Hughes, J. R., Higgs, D. R. and Tufarelli, C. (2017) 'DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease', *Proc Natl Acad Sci U S A,* 114(36), pp. E7526-E7535.

Jia, R. Z., Zhang, X., Hu, P., Liu, X. M., Hua, X. D., Wang, X. and Ding, H. J. (2012) 'Screening for differential methylation status in human placenta in preeclampsia using a CpG island plus promoter microarray', *Int J Mol Med,* 30(1), pp. 133-41.

Jia, Y., Li, T., Huang, X., Xu, X., Zhou, X., Jia, L., Zhu, J., Xie, D., Wang, K., Zhou, Q., Jin, L., Zhang, J. and Duan, T. (2017) 'Dysregulated DNA Methyltransferase 3A Upregulates IGFBP5 to Suppress Trophoblast Cell Migration and Invasion in Preeclampsia', *Hypertension,* 69(2), pp. 356-366.

Jiang, P., Chan, C. W., Chan, K. C., Cheng, S. H., Wong, J., Wong, V. W., Wong, G. L., Chan, S. L., Mok, T. S., Chan, H. L., Lai, P. B., Chiu, R. W. and Lo, Y. M. (2015) 'Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients', *Proc Natl Acad Sci U S A,* 112(11), pp. E1317-25.

Jim, B. and Karumanchi, S. A. (2017) 'Preeclampsia: Pathogenesis, Prevention, and Long-Term Complications', *Semin Nephrol,* 37(4), pp. 386-397.

Jin, Z. and Liu, Y. (2018) 'DNA methylation in human diseases', *Genes Dis,* 5(1), pp. 1-8.

Jones, P. A. (2012) 'Functions of DNA methylation: islands, start sites, gene bodies and beyond', *Nat Rev Genet,* 13(7), pp. 484-92.

Joo, J. E., Hiden, U., Lassance, L., Gordon, L., Martino, D. J., Desoye, G. and Saffery, R. (2013) 'Variable promoter methylation contributes to differential expression of key genes in human placenta-derived venous and arterial endothelial cells', *BMC Genomics,* 14, pp. 475.

Kalafat, E., Sukur, Y. E., Abdi, A., Thilaganathan, B. and Khalil, A. (2018) 'Metformin for prevention of hypertensive disorders of pregnancy in women with gestational diabetes or obesity: systematic review and meta-analysis of randomized trials', *Ultrasound Obstet Gynecol,* 52(6), pp. 706-714.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J., Amin, V., Whitaker, J. W., Schultz, M. D., Ward, L. D., Sarkar, A., Quon, G., Sandstrom, R. S., Eaton, M. L., Wu, Y. C., Pfenning, A. R., Wang, X., Claussnitzer, M., Liu, Y., Coarfa, C., Harris, R. A., Shoresh, N., Epstein, C. B., Gjoneska, E., Leung, D., Xie, W., Hawkins, R. D., Lister, R., Hong, C., Gascard, P., Mungall, A. J., Moore, R., Chuah, E., Tam, A., Canfield, T. K., Hansen, R. S., Kaul, R., Sabo, P. J., Bansal, M. S., Carles, A., Dixon, J. R., Farh, K. H., Feizi, S., Karlic, R., Kim, A. R., Kulkarni, A., Li, D., Lowdon, R., Elliott, G., Mercer, T. R., Neph, S. J., Onuchic, V., Polak, P., Rajagopal, N., Ray, P., Sallari, R. C., Siebenthall, K. T., Sinnott-Armstrong, N. A., Stevens, M., Thurman, R. E., Wu, J., Zhang, B., Zhou, X., Beaudet, A. E., Boyer, L. A., De Jager, P. L., Farnham, P. J., Fisher, S. J., Haussler, D., Jones, S. J., Li, W., Marra, M. A., McManus, M. T., Sunyaev, S., Thomson, J. A., Tlsty, T. D., Tsai, L. H., Wang, W., Waterland, R. A., Zhang, M. Q., Chadwick, L. H., Bernstein, B. E., Costello, J. F., Ecker, J. R., Hirst, M., Meissner, A., Milosavljevic, A., Ren, B., Stamatoyannopoulos, J. A., Wang, T., Kellis, M. and Consortium, R. E. (2015) 'Integrative analysis of 111 reference human epigenomes', *Nature,* 518(7539), pp. 317-30.

Kurdyukov, S. and Bullock, M. (2016) 'DNA Methylation Analysis: Choosing the Right Method', *Biology (Basel),* 5(1).

Kustanovich, A., Schwartz, R., Peretz, T. and Grinshpun, A. (2019) 'Life and death of circulating cell-free DNA', *Cancer Biol Ther*, pp. 1-11.

Lau, T. W., Leung, T. N., Chan, L. Y., Lau, T. K., Chan, K. C., Tam, W. H. and Lo, Y. M. (2002) 'Fetal DNA clearance from maternal plasma is impaired in preeclampsia', *Clin Chem,* 48(12), pp. 2141-6.

Leavey, K., Wilson, S. L., Bainbridge, S. A., Robinson, W. P. and Cox, B. J. (2018) 'Epigenetic regulation of placental gene expression in transcriptional subtypes of preeclampsia', *Clin Epigenetics,* 10, pp. 28.

Lehmann-Werman, R., Neiman, D., Zemmour, H., Moss, J., Magenheim, J., Vaknin-Dembinsky, A., Rubertsson, S., Nellgård, B., Blennow, K., Zetterberg, H., Spalding, K., Haller, M. J., Wasserfall, C. H., Schatz, D. A., Greenbaum, C. J., Dorrell, C., Grompe, M., Zick, A., Hubert, A., Maoz, M., Fendrich, V., Bartsch, D. K., Golan, T., Ben Sasson, S. A., Zamir, G., Razin, A., Cedar, H., Shapiro, A. M., Glaser, B., Shemer, R. and Dor, Y. (2016) 'Identification of tissue-specific cell death using methylation patterns of circulating DNA', *Proc Natl Acad Sci U S A,* 113(13), pp. E1826-34.

Leon, S. A., Shapiro, B., Sklaroff, D. M. and Yaros, M. J. (1977) 'Free DNA in the serum of cancer patients and the effect of therapy', *Cancer Res,* 37(3), pp. 646-50.

Li, H. (2011) 'A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data', *Bioinformatics,* 27(21), pp. 2987-93.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics,* 25(14), pp. 1754-60.

Li, J., Huang, Q., Zeng, F., Li, W., He, Z., Chen, W., Zhu, W. and Zhang, B. (2014) 'The prognostic value of global DNA hypomethylation in cancer: a meta-analysis', *PLoS One,* 9(9), pp. e106290.

Li, X., Tjwa, M., Van Hove, I., Enholm, B., Neven, E., Paavonen, K., Jeltsch, M., Juan, T. D., Sievers, R. E., Chorianopoulos, E., Wada, H., Vanwildemeersch, M., Noel, A., Foidart, J. M., Springer, M. L., von Degenfeld, G., Dewerchin, M., Blau, H. M., Alitalo, K., Eriksson, U., Carmeliet, P. and Moons, L. (2008) 'Reevaluation of the role of VEGF-B suggests a restricted role in the revascularization of the ischemic myocardium', *Arterioscler Thromb Vasc Biol,* 28(9), pp. 1614-20.

Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z. and Zhang, B. (2019) 'WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs', *Nucleic Acids Res,* 47(W1), pp. W199-w205.

Lim, Y. C., Li, J., Ni, Y., Liang, Q., Zhang, J., Yeo, G. S. H., Lyu, J., Jin, S. and Ding, C. (2017) 'A complex association between DNA methylation and gene expression in human placenta at first and third trimesters', *PLoS One,* 12(7), pp. e0181155.

Liu, H., Li, Y., Zhang, J., Rao, M., Liang, H. and Liu, G. (2015) 'The defect of both angiogenesis and lymphangiogenesis is involved in preeclampsia', *Placenta,* 36(3), pp. 279-86.

Liu, H., Tang, Y., Liu, X., Zhou, Q., Xiao, X., Lan, F., Li, X., Hu, R., Xiong, Y. and Peng, T. (2014a) '14-3-3 tau (YWHAQ) gene promoter hypermethylation in human placenta of preeclampsia', *Placenta,* 35(12), pp. 981-8.

Liu, Q., Qiao, F. Y., Shi, X. W., Liu, H. Y., Gong, X. and Wu, Y. Y. (2014b) 'Promoter hypomethylation and increased maspin expression in preeclamptic placentas in a Chinese population', *Placenta,* 35(11), pp. 876-82.

Liu, R., Ma, Q., Wen, A., Tian, G., Li, M. and Wang, W. (2017a) 'Increased tissue factor expression and promoter hypomethylation in preeclampsia placentas in a Chinese population', *Pregnancy Hypertens,* 10, pp. 90-95.

Liu, Y. and Ma, Y.(2017b) 'Promoter Methylation Status of WNT2 in Placenta from Patients with Preeclampsia', *Med Sci Monit,* 23, pp. 5294-5301.

Lo, Y. M., Chan, K. C., Sun, H., Chen, E. Z., Jiang, P., Lun, F. M., Zheng, Y. W., Leung, T. Y., Lau, T. K., Cantor, C. R. and Chiu, R. W. (2010) 'Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus', *Science translational medicine,* 2(61), pp. 61ra91.

Lo, Y. M., Corbetta, N., Chamberlain, P. F., Rai, V., Sargent, I. L., Redman, C. W. and Wainscoat, J. S. (1997) 'Presence of fetal DNA in maternal plasma and serum', *Lancet,* 350(9076), pp. 485-487.

Lo, Y. M., Leung, T. N., Tein, M. S., Sargent, I. L., Zhang, J., Lau, T. K., Haines, C. J. and Redman, C. W. (1999) 'Quantitative abnormalities of fetal DNA in maternal serum in preeclampsia', *Clin Chem,* 45(2), pp. 184-8.

Lo, Y. M., Tein, M. S., Pang, C. C., Yeung, C. K., Tong, K. L. and Hjelm, N. M. (1998) 'Presence of donor-specific DNA in plasma of kidney and liver-transplant recipients', *Lancet,* 351(9112), pp. 1329-30.

Logue, O. C., Mahdi, F., Chapman, H., George, E. M. and Bidwell, G. L. (2017) 'A Maternally Sequestered, Biopolymer-Stabilized Vascular Endothelial Growth Factor (VEGF) Chimera for Treatment of Preeclampsia', *J Am Heart Assoc,* 6(12).

Lokki, A. I., Heikkinen-Eloranta, J. K. and Laivuori, H. (2018) 'The Immunogenetic Conundrum of Preeclampsia', *Front Immunol,* 9, pp. 2630.

Lu, L., Hou, Z., Li, L., Yang, Y., Wang, X., Zhang, B., Ren, M., Zhao, D., Miao, Z., Yu, L. and Yao, Y. (2014) 'Methylation pattern of H19 exon 1 is closely related to preeclampsia and trophoblast abnormalities', *Int J Mol Med,* 34(3), pp. 765-71.

Lun, F. M., Chiu, R. W., Sun, K., Leung, T. Y., Jiang, P., Chan, K. C., Sun, H. and Lo, Y. M. (2013) 'Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA', *Clin Chem,* 59(11), pp. 1583-94.

Luo, C., Hajkova, P. and Ecker, J. R. (2018) 'Dynamic DNA methylation: In the right place at the right time', *Science,* 361(6409), pp. 1336-1340.

Lykke, J. A., Langhoff-Roos, J., Sibai, B. M., Funai, E. F., Triche, E. W. and Paidas, M. J. (2009) 'Hypertensive pregnancy disorders and subsequent cardiovascular morbidity and type 2 diabetes mellitus in the mother', *Hypertension,* 53(6), pp. 944-51.

Ma, M., Zhou, Q. J., Xiong, Y., Li, B. and Li, X. T. (2018) 'Preeclampsia is associated with hypermethylation of IGF-1 promoter mediated by DNMT1', *Am J Transl Res,* 10(1), pp. 16-39.

Ma, X., Li, J., Brost, B., Cheng, W. and Jiang, S. W. (2014) 'Decreased expression and DNA methylation levels of GATAD1 in preeclamptic placentas', *Cell Signal,* 26(5), pp. 959-67.

Mandel, P. and Metais, P. (1948) 'Les acides nucléiques du plasma sanguin chez l'homme.', *C R Seances Soc Biol Fil,* 142(3-4), pp. 241-3.

Mano, Y., Kotani, T., Shibata, K., Matsumura, H., Tsuda, H., Sumigama, S., Yamamoto, E., Iwase, A., Senga, T. and Kikkawa, F. (2011) 'The loss of endoglin promotes the invasion of extravillous trophoblasts', *Endocrinology,* 152(11), pp. 4386-94.

Martin, A., Krishna, I., Badell, M., Martina, B. and Samuel, A. (2014) 'Can the quantity of cell-free fetal DNA predict preeclampsia: a systematic review', *Prenat Diagn,* 34(7), pp. 685-91.

Martin, E., Ray, P. D., Smeester, L., Grace, M. R., Boggess, K. and Fry, R. C. (2015) 'Epigenetics and Preeclampsia: Defining Functional Epimutations in the Preeclamptic Placenta Related to the TGF-β Pathway', *PLoS One,* 10(10), pp. e0141294.

Maynard, S. E., Min, J. Y., Merchan, J., Lim, K. H., Li, J., Mondal, S., Libermann, T. A., Morgan, J. P., Sellke, F. W., Stillman, I. E., Epstein, F. H., Sukhatme, V. P. and Karumanchi, S. A. (2003) 'Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia', *J Clin Invest,* 111(5), pp. 649-58.

Moffett-King, A. (2002) 'Natural killer cells and pregnancy', *Nat Rev Immunol,* 2(9), pp. 656-63.

Mohn, F., Weber, M., Rebhan, M., Roloff, T. C., Richter, J., Stadler, M. B., Bibel, M. and Schübeler, D. (2008) 'Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors', *Mol Cell,* 30(6), pp. 755-66.

Molnár, M., Sütö, T., Tóth, T. and Hertelendy, F. (1994) 'Prolonged blockade of nitric oxide synthesis in gravid rats produces sustained hypertension, proteinuria, thrombocytopenia, and intrauterine growth retardation', *Am J Obstet Gynecol,* 170(5 Pt 1), pp. 1458-66.

Moore, L. D., Le, T. and Fan, G. (2013) 'DNA methylation and its basic function', *Neuropsychopharmacology,* 38(1), pp. 23-38.

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., Houstis, N., Daly, M. J., Patterson, N., Mesirov, J. P., Golub, T. R., Tamayo, P., Spiegelman, B., Lander, E. S., Hirschhorn, J. N., Altshuler, D. and Groop, L. C. (2003) 'PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes', *Nat Genet,* 34(3), pp. 267-73.

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

Moss, J., Magenheim, J., Neiman, D., Zemmour, H., Loyfer, N., Korach, A., Samet, Y., Maoz, M., Druid, H., Arner, P., Fu, K. Y., Kiss, E., Spalding, K. L., Landesberg, G., Zick, A., Grinshpun, A., Shapiro, A. M. J., Grompe, M., Wittenberg, A. D., Glaser, B., Shemer, R., Kaplan, T. and Dor, Y. (2018) 'Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease', *Nat Commun,* 9(1), pp. 5068.

Mouliere, F., Chandrananda, D., Piskorz, A. M., Moore, E. K., Morris, J., Ahlborn, L. B., Mair, R., Goranova, T., Marass, F., Heider, K., Wan, J. C. M., Supernat, A., Hudecova, I., Gounaris, I., Ros, S., Jimenez-Linan, M., Garcia-Corbacho, J., Patel, K., Østrup, O., Murphy, S., Eldridge, M. D., Gale, D., Stewart, G. D., Burge, J., Cooper, W. N., van der Heijden, M. S., Massie, C. E., Watts, C., Corrie, P., Pacey, S., Brindle, K. M., Baird, R. D., Mau-Sørensen, M., Parkinson, C. A., Smith, C. G., Brenton, J. D. and Rosenfeld, N. (2018) 'Enhanced detection of circulating tumor DNA by fragment size analysis', *Sci Transl Med,* 10(466).

Mousa, A. A., Archer, K. J., Cappello, R., Estrada-Gutierrez, G., Isaacs, C. R., Strauss, J. F. and Walsh, S. W. (2012a) 'DNA methylation is altered in maternal blood vessels of women with preeclampsia', *Reprod Sci,* 19(12), pp. 1332-42.

Mousa, A. A., Cappello, R. E., Estrada-Gutierrez, G., Shukla, J., Romero, R., Strauss, J. F. and Walsh, S. W. (2012b) 'Preeclampsia is associated with alterations in DNA methylation of genes involved in collagen metabolism', *Am J Pathol,* 181(4), pp. 1455-63.

Mousa, A. A., Strauss, J. F. and Walsh, S. W. (2012c) 'Reduced methylation of the thromboxane synthase gene is correlated with its increased vascular expression in preeclampsia', *Hypertension,* 59(6), pp. 1249-55.

Nair, S. S., Coolen, M. W., Stirzaker, C., Song, J. Z., Statham, A. L., Strbenac, D., Robinson, M. D. and Clark, S. J. (2011) 'Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias', *Epigenetics,* 6(1), pp. 34-44.

Neidhart, M. (2015) *DNA Methylation and Complex Human Disease.*

Novakovic, B., Rakyan, V., Ng, H. K., Manuelpillai, U., Dewi, C., Wong, N. C., Morley, R., Down, T., Beck, S., Craig, J. M. and Saffery, R. (2008) 'Specific tumour-associated methylation in normal human term placenta and first-trimester cytotrophoblasts', *Mol Hum Reprod,* 14(9), pp. 547-54.

Paauw, N. D., Terstappen, F., Ganzevoort, W., Joles, J. A., Gremmels, H. and Lely, A. T. (2017) 'Sildenafil During Pregnancy: A Preclinical Meta-Analysis on Fetal Growth and Maternal Blood Pressure', *Hypertension,* 70(5), pp. 998-1006.

Pandey, A. K., Singhi, E. K., Arroyo, J. P., Ikizler, T. A., Gould, E. R., Brown, J., Beckman, J. A., Harrison, D. G. and Moslehi, J. (2018) 'Mechanisms of VEGF (Vascular Endothelial Growth Factor) Inhibitor-Associated Hypertension and Vascular Disease', *Hypertension,* 71(2), pp. e1-e8.

Phipps, E. A., Thadhani, R., Benzing, T. and Karumanchi, S. A. (2019) 'Pre-eclampsia: pathogenesis, novel diagnostics and therapies', *Nat Rev Nephrol*.

Powe, C. E., Levine, R. J. and Karumanchi, S. A. (2011) 'Preeclampsia, a disease of the maternal endothelium: the role of antiangiogenic factors and implications for later cardiovascular disease', *Circulation,* 123(24), pp. 2856-69.

Probst, A. V., Dunleavy, E. and Almouzni, G. (2009) 'Epigenetic inheritance during the cell cycle', *Nat Rev Mol Cell Biol,* 10(3), pp. 192-206.

Purwosunu, Y., Sekizawa, A., Yoshimura, S., Farina, A., Wibowo, N., Nakamura, M., Shimizu, H. and Okai, T. (2009) 'Expression of angiogenesis-related genes in the cellular component of the blood of preeclamptic women', *Reprod Sci,* 16(9), pp. 857-64.

Rahat, B., Najar, R. A., Hamid, A., Bagga, R. and Kaur, J. (2017) 'The role of aberrant methylation of trophoblastic stem cell origin in the pathogenesis and diagnosis of placental disorders', *Prenat Diagn,* 37(2), pp. 133-143.

Rahat, B., Sharma, R., Bagga, R., Hamid, A. and Kaur, J. (2016a) 'Imbalance between matrix metalloproteinases and their tissue inhibitors in preeclampsia and gestational trophoblastic diseases', *Reproduction,* 152(1), pp. 11-22.

Rahat, B., Thakur, S., Hamid, A., Bagga, R. and Kaur, J. (2016b) 'Association of aberrant methylation at promoter regions of tumor suppressor genes with placental pathologies', *Epigenomics,* 8(6), pp. 767-87.

Rajakumar, A., Brandon, H. M., Daftary, A., Ness, R. and Conrad, K. P. (2004) 'Evidence for the functional activity of hypoxia-inducible transcription factors overexpressed in preeclamptic placentae', *Placenta,* 25(10), pp. 763-9.

Ramirez, F., Ryan, D. P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dundar, F. and Manke, T. (2016) 'deepTools2: a next generation web server for deep-sequencing data analysis', *Nucleic Acids Res,* 44(W1), pp. W160-5.

Redman, C. W. and Sargent, I. L. (2005) 'Latest advances in understanding preeclampsia', *Science,* 308(5728), pp. 1592-4.

Redman, C. W. and Sargent, I. L. (2010) 'Immunology of pre-eclampsia', *Am J Reprod Immunol,* 63(6), pp. 534-43.

Robertson, K. D. (2005) 'DNA methylation and human disease', *Nat Rev Genet,* 6(8), pp. 597-610.

Rolnik, D. L., da Silva Costa, F., Lee, T. J., Schmid, M. and McLennan, A. C. (2018) 'Association between fetal fraction on cell-free DNA testing and first-trimester markers for pre-eclampsia', *Ultrasound Obstet Gynecol,* 52(6), pp. 722-727.

Rolnik, D. L., O'Gorman, N., Fiolna, M., van den Boom, D., Nicolaides, K. H. and Poon, L. C. (2015) 'Maternal plasma cell-free DNA in the prediction of pre-eclampsia', *Ultrasound Obstet Gynecol,* 45(1), pp. 106-11.

Sari, I., Pinarbasi, H., Pinarbasi, E. and Yildiz, C. (2017) 'Association between the soluble epoxide hydrolase gene and preeclampsia', *Hypertens Pregnancy,* 36(4), pp. 315-325.

Schroeder, D. I., Blair, J. D., Lott, P., Yu, H. O., Hong, D., Crary, F., Ashwood, P., Walker, C., Korf, I., Robinson, W. P. and LaSalle, J. M. (2013) 'The human placenta methylome', *Proc Natl Acad Sci U S A,* 110(15), pp. 6037-42.

Serre, D., Lee, B. H. and Ting, A. H. (2010) 'MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome', *Nucleic Acids Res,* 38(2), pp. 391-9.

Shimanuki, Y., Mitomi, H., Fukumura, Y., Makino, S., Itakura, A., Yao, T. and Takeda, S. (2015) 'Alteration of Delta-like ligand 1 and Notch 1 receptor in various placental disorders with special reference to early onset preeclampsia', *Hum Pathol,* 46(8), pp. 1129-37.

Shirodkar, A. V., St Bernard, R., Gavryushova, A., Kop, A., Knight, B. J., Yan, M. S., Man, H. S., Sud, M., Hebbel, R. P., Oettgen, P., Aird, W. C. and Marsden, P. A. (2013) 'A mechanistic role for DNA methylation in endothelial cell (EC)-enriched gene expression: relationship with DNA replication timing', *Blood,* 121(17), pp. 3531-40.

Sitras, V., Paulssen, R. H., Grønaas, H., Leirvik, J., Hanssen, T. A., Vårtun, A. and Acharya, G. (2009) 'Differential placental gene expression in severe preeclampsia', *Placenta,* 30(5), pp. 424-33.

Smith, Z. D. and Meissner, A. (2013) 'DNA methylation: roles in mammalian development', *Nat Rev Genet,* 14(3), pp. 204-20.

Snyder, M. W., Kircher, M., Hill, A. J., Daza, R. M. and Shendure, J. (2016) 'Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin', *Cell,* 164(1-2), pp. 57-68.

Sorenson, G. D., Pribish, D. M., Valone, F. H., Memoli, V. A., Bzik, D. J. and Yao, S. L. (1994) 'Soluble normal and mutated DNA sequences from single-copy genes in human blood', *Cancer Epidemiol Biomarkers Prev,* 3(1), pp. 67-71.

Stroun, M., Anker, P., Maurice, P., Lyautey, J., Lederrey, C. and Beljanski, M. (1989) 'Neoplastic characteristics of the DNA found in the plasma of cancer patients', *Oncology,* 46(5), pp. 318-22.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. and Mesirov, J. P. (2005) 'Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles', *Proc Natl Acad Sci U S A,* 102(43), pp. 15545-50.

Sun, K., Jiang, P., Chan, K. C., Wong, J., Cheng, Y. K., Liang, R. H., Chan, W. K., Ma, E. S., Chan, S. L., Cheng, S. H., Chan, R. W., Tong, Y. K., Ng, S. S., Wong, R. S., Hui, D. S., Leung, T. N., Leung, T. Y., Lai, P. B., Chiu, R. W. and Lo, Y. M. (2015) 'Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments', *Proc Natl Acad Sci U S A,* 112(40), pp. E5503-12.

Sun, K., Jiang, P., Wong, A. I. C., Cheng, Y. K. Y., Cheng, S. H., Zhang, H., Chan, K. C. A., Leung, T. Y., Chiu, R. W. K. and Lo, Y. M. D. (2018a) 'Size-tagged preferred ends in maternal plasma DNA shed light on the production mechanism and show utility in noninvasive prenatal testing', *Proc Natl Acad Sci U S A,* 115(22), pp. E5106-E5114.

Sun, K., Lun, F. M. F., Leung, T. Y., Chiu, R. W. K., Lo, Y. M. D. and Sun, H. (2018b) 'Noninvasive reconstruction of placental methylome from maternal plasma DNA: Potential for prenatal testing and monitoring', *Prenat Diagn,* 38(3), pp. 196-203.

Sun, X. J., Wang, M. C., Zhang, F. H. and Kong, X. (2018c) 'An integrated analysis of genome-wide DNA methylation and gene expression data in hepatocellular carcinoma', *FEBS Open Bio,* 8(7), pp. 1093-1103.

Suzuki, M., Maekawa, R., Patterson, N. E., Reynolds, D. M., Calder, B. R., Reznik, S. E., Heo, H. J., Einstein, F. H. and Greally, J. M. (2016) 'Amnion as a surrogate tissue reporter of the effects of maternal preeclampsia on the fetus', *Clin Epigenetics,* 8, pp. 67.

Taiwo, O., Wilson, G. A., Morris, T., Seisenberger, S., Reik, W., Pearce, D., Beck, S. and Butcher, L. M. (2012) 'Methylome analysis using MeDIP-seq with low DNA concentrations', *Nat Protoc,* 7(4), pp. 617-36.

Tan, E. M., Schur, P. H., Carr, R. I. and Kunkel, H. G. (1966) 'Deoxybonucleic acid (DNA) and antibodies to DNA in the serum of patients with systemic lupus erythematosus', *J Clin Invest,* 45(11), pp. 1732-40.

Tang, Y., Liu, H., Li, H., Peng, T., Gu, W. and Li, X. (2015) 'Hypermethylation of the HLA-G promoter is associated with preeclampsia', *Mol Hum Reprod,* 21(9), pp. 736-44.

Than, N. G., Romero, R., Xu, Y., Erez, O., Xu, Z., Bhatti, G., Leavitt, R., Chung, T. H., El-Azzamy, H., LaJeunesse, C., Wang, B., Balogh, A., Szalai, G., Land, S., Dong, Z., Hassan, S. S., Chaiworapongsa, T., Krispin, M., Kim, C. J., Tarca, A. L., Papp, Z. and Bohn, H. (2014) 'Evolutionary origins of the placental expression of chromosome 19 cluster galectins and their complex dysregulation in preeclampsia', *Placenta,* 35(11), pp. 855-65.

Ulz, P., Thallinger, G. G., Auer, M., Graf, R., Kashofer, K., Jahn, S. W., Abete, L., Pristauz, G., Petru, E., Geigl, J. B., Heitzer, E. and Speicher, M. R. (2016) 'Inferring expressed genes by whole-genome sequencing of plasma DNA', *Nat Genet,* 48(10), pp. 1273-8.

van den Berg, C. B., Chaves, I., Herzog, E. M., Willemsen, S. P., van der Horst, G. T. J. and Steegers-Theunissen, R. P. M. (2017) 'Early- and late-onset preeclampsia and the DNA methylation of circadian clock and clock-controlled genes in placental and newborn tissues', *Chronobiol Int,* 34(7), pp. 921-932.

Vasioukhin, V., Anker, P., Maurice, P., Lyautey, J., Lederrey, C. and Stroun, M. (1994) 'Point mutations of the N-ras gene in the blood plasma DNA of patients with myelodysplastic syndrome or acute myelogenous leukaemia', *Br J Haematol,* 86(4), pp. 774-9.

Venkatesha, S., Toporsian, M., Lam, C., Hanai, J., Mammoto, T., Kim, Y. M., Bdolah, Y., Lim, K. H., Yuan, H. T., Libermann, T. A., Stillman, I. E., Roberts, D., D'Amore, P. A., Epstein, F. H., Sellke, F. W., Romero, R., Sukhatme, V. P., Letarte, M. and Karumanchi, S. A. (2006) 'Soluble endoglin contributes to the pathogenesis of preeclampsia', *Nat Med,* 12(6), pp. 642-9.

Volik, S., Alcaide, M., Morin, R. D. and Collins, C. (2016) 'Cell-free DNA (cfDNA): Clinical Significance and Utility in Cancer Shaped By Emerging Technologies', *Mol Cancer Res,* 14(10), pp. 898-908.

von Meyenn, F. and Reik, W. (2015) 'Forget the Parents: Epigenetic Reprogramming in Human Germ Cells', *Cell,* 161(6), pp. 1248-51.

Walker, D. L., Bhagwate, A. V., Baheti, S., Smalley, R. L., Hilker, C. A., Sun, Z. and Cunningham, J. M. (2015) 'DNA methylation profiling: comparison of genome-wide sequencing methods and the Infinium Human Methylation 450 Bead Chip', *Epigenomics,* 7(8), pp. 1287-302.

Waller, J. P., Burke, S. P., Logue, O. C. and and Bidwell, G. L. 'A Biopolymer-Fused Form of VEGF-B for Treatment of Preeclampsia', *Experimental Biology 2019 Meeting*.

Wan, J. C. M., Massie, C., Garcia-Corbacho, J., Mouliere, F., Brenton, J. D., Caldas, C., Pacey, S., Baird, R. and Rosenfeld, N. (2017) 'Liquid biopsies come of age: towards implementation of circulating tumour DNA', *Nat Rev Cancer,* 17(4), pp. 223-238.

Wang, A., Rana, S. and Karumanchi, S. A. (2009) 'Preeclampsia: the role of angiogenic factors in its pathogenesis', *Physiology (Bethesda),* 24, pp. 147-58.

Wang, X., Zhang, Z., Zeng, X., Wang, J., Zhang, L., Song, W. and Shi, Y. (2018) 'Wnt/β-catenin signaling pathway in severe preeclampsia', *J Mol Histol,* 49(3), pp. 317-327.

Wang, Z., Lu, S., Liu, C., Zhao, B., Pei, K., Tian, L. and Ma, X. (2010) 'Expressional and epigenetic alterations of placental matrix metalloproteinase 9 in preeclampsia', *Gynecol Endocrinol,* 26(2), pp. 96-102.

White, W. M., Brost, B., Sun, Z., Rose, C., Craici, I., Wagner, S. J., Turner, S. T. and Garovic, V. D. (2013) 'Genome-wide methylation profiling demonstrates hypermethylation in maternal leukocyte DNA in preeclamptic compared to normotensive pregnancies', *Hypertens Pregnancy,* 32(3), pp. 257-69.

White, W. M., Sun, Z., Borowski, K. S., Brost, B. C., Davies, N. P., Rose, C. H. and Garovic, V. D. (2016) 'Preeclampsia/Eclampsia candidate genes show altered methylation in maternal leukocytes of preeclamptic women at the time of delivery', *Hypertens Pregnancy,* 35(3), pp. 394-404.

Wilson, S. L., Blair, J. D., Hogg, K., Langlois, S., von Dadelszen, P. and Robinson, W. P. (2015) 'Placental DNA methylation at term reflects maternal serum levels of INHA and FN1, but not PAPPA, early in pregnancy', *BMC Med Genet,* 16, pp. 111.

Wright, M. L., Anderson, C. M., Uthus, E. O. and Ohm, J. E. (2012) 'Validation of DNA methylation patterns: potential biomarker for heritable risk of preeclampsia', *West J Nurs Res,* 34(8), pp. 1074-5.

Xiang, Y., Zhang, X., Li, Q., Xu, J., Zhou, X., Wang, T., Xing, Q., Liu, Y., Wang, L., He, L. and Zhao, X. (2013) 'Promoter hypomethylation of TIMP3 is associated with pre-eclampsia in a Chinese population', *Mol Hum Reprod,* 19(3), pp. 153-9.

Xiao, X., Tao, X., Wang, Y., Zhu, L., Ye, Y., Liu, H., Zhou, Q., Li, X. and Xiong, Y. (2017) 'Hypomethylation of tissue factor pathway inhibitor 2 in human placenta of preeclampsia', *Thromb Res,* 152, pp. 7-13.

Xuan, J., Jing, Z., Yuanfang, Z., Xiaoju, H., Pei, L., Guiyin, J. and Yu, Z. (2016) 'Comprehensive analysis of DNA methylation and gene expression of placental tissue in preeclampsia patients', *Hypertens Pregnancy,* 35(1), pp. 129-38.

Xuan Lin, Q. X., Sian, S., An, O., Thieffry, D., Jha, S. and Benoukraf, T. (2019) 'MethMotif: an integrative cell specific database of transcription factor binding motifs coupled with DNA methylation profiles', *Nucleic Acids Res,* 47(D1), pp. D145-D154.

Yang, X., Han, H., De Carvalho, D. D., Lay, F. D., Jones, P. A. and Liang, G. (2014) 'Gene body methylation can alter gene expression and is a therapeutic target in cancer', *Cancer Cell,* 26(4), pp. 577-90.

Ye, W., Shen, L., Xiong, Y., Zhou, Y., Gu, H. and Yang, Z. (2016) 'Preeclampsia is Associated with Decreased Methylation of the GNA12 Promoter', *Ann Hum Genet,* 80(1), pp. 7-10.

Yeung, K. R., Chiu, C. L., Pidsley, R., Makris, A., Hennessy, A. and Lind, J. M. (2016) 'DNA methylation profiles in preeclampsia and healthy control placentas', *Am J Physiol Heart Circ Physiol,* 310(10), pp. H1295-303.

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P. K., Kivioja, T., Dave, K., Zhong, F., Nitta, K. R., Taipale, M., Popov, A., Ginno, P. A., Domcke, S., Yan, J., Schübeler, D., Vinson, C. and Taipale, J. (2017) 'Impact of cytosine methylation on DNA binding specificities of human transcription factors', *Science,* 356(6337).

Yong, W. S., Hsu, F. M. and Chen, P. Y. (2016) 'Profiling genome-wide DNA methylation', *Epigenetics Chromatin,* 9, pp. 26.

Young, B. C., Levine, R. J. and Karumanchi, S. A. (2010) 'Pathogenesis of preeclampsia', *Annu Rev Pathol,* 5, pp. 173-92.

Yu, S. C., Chan, K. C., Zheng, Y. W., Jiang, P., Liao, G. J., Sun, H., Akolekar, R., Leung, T. Y., Go, A. T., van Vugt, J. M., Minekawa, R., Oudejans, C. B., Nicolaides, K. H., Chiu, R. W. and Lo, Y. M. (2014) 'Size-based molecular diagnostics using plasma DNA for noninvasive prenatal testing', *Proc Natl Acad Sci U S A,* 111(23), pp. 8583-8.

Yu, S. C., Lee, S. W., Jiang, P., Leung, T. Y., Chan, K. C., Chiu, R. W. and Lo, Y. M. (2013) 'High-resolution profiling of fetal DNA clearance from maternal plasma by massively parallel sequencing', *Clinical chemistry,* 59(8), pp. 1228-1237.

Yuen, R. K., Avila, L., Peñaherrera, M. S., von Dadelszen, P., Lefebvre, L., Kobor, M. S. and Robinson, W. P. (2009) 'Human placental-specific epipolymorphism and its association with adverse pregnancy outcomes', *PLoS One,* 4(10), pp. e7389.

Yuen, R. K., Peñaherrera, M. S., von Dadelszen, P., McFadden, D. E. and Robinson, W. P. (2010) 'DNA methylation profiling of human placentas reveals promoter hypomethylation of multiple genes in early-onset preeclampsia', *Eur J Hum Genet,* 18(9), pp. 1006-12.

Zeng, F., Zhu, S., Wong, M. C., Yang, Z., Tang, J., Li, K. and Su, X. (2016) 'Associations between nitric oxide synthase 3 gene polymorphisms and preeclampsia risk: a meta-analysis', *Sci Rep,* 6, pp. 23407.

Zhang, L., Leng, M., Li, Y., Yuan, Y., Yang, B., Yuan, E., Shi, W., Yan, S. and Cui, S. (2019) 'Altered DNA methylation and transcription of WNT2 and DKK1 genes in placentas associated with early-onset preeclampsia', *Clin Chim Acta,* 490, pp. 154-160.

Zhang, R., Nakahira, K., Guo, X., Choi, A. M. and Gu, Z. (2016) 'Very Short Mitochondrial DNA Fragments and Heteroplasmy in Human Plasma', *Sci Rep,* 6, pp. 36097.

Zhao, M., Li, L., Yang, X., Cui, J. and Li, H. (2017) 'FN1, FOS, and ITGA5 induce preeclampsia: Abnormal expression and methylation', *Hypertens Pregnancy,* 36(4), pp. 302-309.

Zheng, Y. W., Chan, K. C., Sun, H., Jiang, P., Su, X., Chen, E. Z., Lun, F. M., Hung, E. C., Lee, V., Wong, J., Lai, P. B., Li, C. K., Chiu, R. W. and Lo, Y. M. (2012) 'Nonhematopoietically derived DNA is shorter than hematopoietically derived DNA in plasma: a transplantation model', *Clin Chem,* 58(3), pp. 549-58.

Zhu, H., Wang, G. and Qian, J. (2016) 'Transcription factors as readers and effectors of DNA methylation', *Nat Rev Genet,* 17(9), pp. 551-65.

Zhu, L., Lv, R., Kong, L., Cheng, H., Lan, F. and Li, X. (2015) 'Genome-Wide Mapping of 5mC and 5hmC Identified Differentially Modified Genomic Regions in Late-Onset Severe Preeclampsia: A Pilot Study', *PLoS One,* 10(7), pp. e0134119.

Zhu, L., Lv, R., Kong, L. and Li, X. (2018) 'Reduced methylation downregulates CD39/ENTPD1 and ZDHHC14 to suppress trophoblast cell proliferation and invasion: Implications in preeclampsia', *Pregnancy Hypertens,* 14, pp. 59-67.

Zhuang, X. W., Li, J., Brost, B. C., Xia, X. Y., Chen, H. B., Wang, C. X. and Jiang, S. W. (2014) 'Decreased expression and altered methylation of syncytin-1 gene in human placentas associated with preeclampsia', *Curr Pharm Des,* 20(11), pp. 1796-802.

Øster, B., Linnet, L., Christensen, L. L., Thorsen, K., Ongen, H., Dermitzakis, E. T., Sandoval, J., Moran, S., Esteller, M., Hansen, T. F., Lamy, P., Laurberg, S., Ørntoft, T. F., Andersen, C. L. and group, C. s. (2013) 'Non-CpG island promoter hypomethylation and miR-149 regulate the expression of SRPX2 in colorectal cancer', *Int J Cancer,* 132(10), pp. 2303-15.

IX. ANNEXES

IX. ANNEXES

ANNEX 1. Run 1- Transcription Factors Binding Sites Over-represented

- **Hypomethylated Regions**

TFBS
ARID3A
ATF1
BATF
BCL11A
CEBPB
EP300
FOS
FOXA1
GATA2
GATA3
GRp20

TFBS
IKZF1
IRF3
IRF4
JUN
KAP1
MAFF
MAFK
MBD4
MEF2C
MTA3
NFE2

TFBS
NFIC
PRDM1
SMARCC2
SPI1
STAT2
STAT3
TAL1
TCF7L2
TRIM28
ZNF217
ZNF274

- **Hypermethylated Regions**

TFBS
BACH1
CCNT2
CTBP2
CTCFL
E2F6
EGR1
EZH2
GABPA
GTF3C2
HDAC1

TFBS
HDAC6
HMGN3
KDM5B
MAZ
NR2C2
PHF8
POLR2A
RBBP5
RPC155
SAP30

TFBS
SMARCB1
SP4
SUZ12
TAF7
TCF3
UBTF
ZBTB33
ZBTB7A
ZEB1
ZNF263

ANNEX 2. Run 2- Transcription Factors Binding Sites Over-represented

- **Hypomethylated Regions**

TFBS
ARID3A
ATF1
ATF2
BATF
BCL11A
BRF2
CEBPB
ELK4
EP300
FOS
FOXA1
FOXA2
FOXM1
GATA1
GATA2
GATA3
GRp20

TFBS
GTF2B
IKZF1
IRF3
IRF4
JUN
KAP1
KDM5A
MAFF
MAFK
MBD4
MEF2A
MEF2C
MTA3
NANOG
NFIC
PML
POU5F1

TFBS
PPARGC1A
PRDM1
RDBP
RELA
RFX5
RUNX3
SIRT6
SIX5
SPI1
STAT1
STAT2
STAT3
TAL1
TRIM28
WRNIP1
ZNF274
ZZZ3

- **Hypermethylated Regions**

TFBS
CTBP2
CTCFL
E2F6
EGR1
ESRRA
EZH2
GABPA
GTF3C2

TFBS
HMGN3
POLR2A
RPC155
SMARCB1
SUZ12
TCF3
ZBTB7A

ANNEX 3. List of genes previously related to preeclampsia present among our DMRS.

Gene name	Gene description
ADCY6	adenylate cyclase 6
ADRA2B	adrenoceptor alpha 2B
AOC3	amine oxidase, copper containing 3
AP2A1	adaptor-related protein complex 2, alpha 1 subunit
APBB3	amyloid beta (A4) precursor protein-binding, family B, member 3
ARHGEF10	Rho guanine nucleotide exchange factor (GEF) 10
ARHGEF15	Rho guanine nucleotide exchange factor (GEF) 15
ARHGEF19	Rho guanine nucleotide exchange factor (GEF) 19
ATP2A3	ATPase, Ca++ transporting, ubiquitous
BACE1	beta-site APP-cleaving enzyme 1
BAZ1A	bromodomain adjacent to zinc finger domain, 1A
CBX7	chromobox homolog 7
CD81	CD81 molecule
COL1A1	collagen, type I, alpha 1
COMT	catechol-O-methyltransferase
CORO2A	coronin, actin binding protein, 2A
CYB5R3	cytochrome b5 reductase 3
DENND2D	DENN/MADD domain containing 2D
DNMT3A	DNA (cytosine-5-)-methyltransferase 3 alpha
DOCK6	dedicator of cytokinesis 6
EARS2	glutamyl-tRNA synthetase 2, mitochondrial
ENG	endoglin
EPHB4	EPH receptor B4
F10	coagulation factor X
FAM69B	family with sequence similarity 69, member B
FAXDC2	fatty acid hydroxylase domain containing 2
GCNT1	glucosaminyl (N-acetyl) transferase 1, core 2
GDA	guanine deaminase
GDF15	growth differentiation factor 15
GUK1	guanylate kinase 1
HAGH	hydroxyacylglutathione hydrolase
HIGD1B	HIG1 hypoxia inducible domain family, member 1B
HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase
HOXA4	homeobox A4
KCTD15	potassium channel tetramerization domain containing 15
KIAA1598	Shootin 1

ANNEX 3. List of genes previously related to preeclampsia present among our DMRS. (Continuation)

Gene name	Gene description
KRT23	keratin 23 (histone deacetylase inducible)
LDLR	low density lipoprotein receptor
LIG1	ligase I, DNA, ATP-dependent
LMF1	lipase maturation factor 1
LPHN1	latrophilin 1
LRIG1	leucine-rich repeats and immunoglobulin-like domains 1
LTBP3	latent transforming growth factor beta binding protein 3
MLLT1	myeloid/lymphoid or mixed-lineage leukemia; translocated to, 1
MMP23B	matrix metallopeptidase 23B
MTUS1	microtubule associated tumor suppressor 1
MUS81	MUS81 structure-specific endonuclease subunit
MYO16	myosin XVI
NADSYN1	NAD synthetase 1
NFATC4	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent4
NOS3	nitric oxide synthase 3 (endothelial cell)
NOTCH3	notch 3
NPHP4	nephronophthisis 4
NR1H2	nuclear receptor subfamily 1, group H, member 2
NSMF	NMDA receptor synaptonuclear signaling and neuronal migration factor
NTRK3	neurotrophic tyrosine kinase, receptor, type 3
OSBPL7	oxysterol binding protein-like 7
PLAT	plasminogen activator, tissue
PLCG1	phospholipase C, gamma 1
PLEKHA4	pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 4
PLEKHG4	pleckstrin homology domain containing, family G (with RhoGef domain) member 4
PPP1R12C	protein phosphatase 1, regulatory subunit 12C
PTP4A2	protein tyrosine phosphatase type IVA, member 2
RANBP3	RAN binding protein 3
RASL12	RAS-like, family 12
RBP5	retinol binding protein 5, cellular
RFTN2	raftlin family member 2
RYR1	ryanodine receptor 1 (skeletal)
SHC2	SHC (Src homology 2 domain containing) transforming protein 2

ANNEX 3. List of genes previously related to preeclampsia present among our DMRS. (Continuation)

Gene name	Gene description
SIAH1	siah E3 ubiquitin protein ligase 1
SLC19A1	solute carrier family 19 (folate transporter), member 1
SLC6A1	solute carrier family 6 (neurotransmitter transporter), member 1
SLIT3	slit homolog 3 (Drosophila)
SMOX	spermine oxidase
SULT2B1	sulfotransferase family, cytosolic, 2B, member 1
TAOK2	TAO kinase 2
THBS2	thrombospondin 2
TMEM129	transmembrane protein 129
TMEM204	transmembrane protein 204
TRAPPC10	trafficking protein particle complex 10
TRMU	tRNA 5-methylaminomethyl-2-thiouridylate methyltransferase
UCKL1	uridine-cytidine kinase 1-like 1
USP21	ubiquitin specific peptidase 21
VANGL2	VANGL planar cell polarity protein 2
VEGFB	vascular endothelial growth factor B
WDR27	WD repeat domain 27
ZNF362	zinc finger protein 362
ZNF423	zinc finger protein 423