

Universidad de Granada

Facultad de Farmacia



PROGRAMA DE DOCTORADO EN FARMACIA

TESIS DOCTORAL:

Efecto de la diabetes gestacional en el epigenoma de la madre y su descendencia.

Teresa María Linares Pineda

Granada, 2023

Directores de Tesis:

Dra. Sonsoles Morcillo Espina

Dra. María José Picón César

Dr. Antonio Sánchez Pozo



UNIVERSIDAD
DE GRANADA



Efecto de la diabetes gestacional en el epigenoma de la madre y su descendencia.

Memoria Tesis Doctoral presentada por:

Teresa María Linares Pineda

Para optar al título de:

Doctora internacional por la Universidad de Granada

Directores de tesis:

Dra. Sonsoles Morcillo Espina

Grupo de investigación obesidad, diabetes y sus comorbilidades: prevención y tratamiento, Instituto de Biomedicina de Málaga, Málaga

Dr. Antonio Sánchez Pozo

Departamento de Bioquímica y Biología Molecular 2. Facultad de Farmacia, Granada.

Dra. María José Picón Cesar

Unidad de Gestión Clínica Endocrinología y Nutrición. Hospital Universitario Virgen de la Victoria, Málaga.

Editor: Universidad de Granada. Tesis Doctorales
Autor: Teresa María Linares Pineda
ISBN: 978-84-1195-216-3
URI: <https://hdl.handle.net/10481/89872>

Trabajos publicados para optar al grado de doctor por la Universidad de Granada

Epigenetic marks associated with Gestational Diabetes Mellitus across two time points during pregnancy. Linares-Pineda T, Peña-Montero N, Frago-Bargas N, Gutiérrez-Repiso C, Lima-Rubio F, Suarez-Arana M, Sánchez-Pozo A, Tinahones FJ, Molina-Vega M, Picón-César MJ, Sommer C, Morcillo S. Clin Epigenetics. 2023 Jul 6;15(1):110. doi: 10.1186/s13148-023-01523-8.

Índices de calidad (Journal Citation Reports):

- Impact factor 2021/2022: 7.28/5.7
- Cuartil 2021/2022: Q1/Q1

Epigenome wide association study in peripheral blood of pregnant women identifies potential metabolic pathways related to gestational diabetes. Linares-Pineda TM, Peña-Montero N, Gutiérrez-Repiso C, Lima-Rubio F, Sánchez-Pozo A, Tinahones FJ, Molina-Vega M, Picón-César MJ, Morcillo S. Epigenetics. 2023 Dec;18(1):2211369. doi: 10.1080/15592294.2023.2211369.

Índices de calidad (Journal Citation Reports):

- Impact factor 2021/2022: 4.8/3.7
- Cuartil 2021/2022: Q1/Q2

Artículo con resultados de la Tesis en revisión:

DNA methylation risk score for type 2 diabetes is associated with gestational diabetes. Teresa M Linares-Pineda, Nicolas Frago-Bargas, María José Picón, Maria Molina-Vega, Anne Karen Jenum, Line Sletner, Sindre Lee-Ødegård, Julia O Opsahl, Gunn-Helen Moen, Elisabeth Qvigstad, Rashmi B Prasad, Kåre I Birkeland, Sonsoles Morcillo, Christine Sommer.

Otras aportaciones científicas durante el periodo de la tesis doctoral:

Epigenetic Marks as Predictors of Metabolic Response to Bariatric Surgery: Validation from an Epigenome Wide Association Study. Carolina Gutiérrez-Repiso, Antonio Cantarero-Cuenca, Andrés González-Jiménez, **Teresa Linares-Pineda**, Nerea Peña-Montero, Luis Ocaña-Wilhelmi, Francisco J Tinahones, Sonsoles Morcillo. *IntJMolSci.* 2023 Sep 30;24(19):14778. doi: 10.3390/ijms241914778. Q1.

Higher β cell death in pregnant women, measured by DNA methylation patterns of cell-free DNA, compared to new-onset type 1 and type 2 diabetes subjects: a cross-sectional study. **Linares-Pineda TM**, Gutiérrez-Repiso C, Peña-Montero N, Molina-Vega M, Rubio FL, Arana MS, Tinahones FJ, Picón-César MJ, Morcillo S. *Diabetol Metab Syndr.* 2023 Jun 1;15(1):115. doi: 10.1186/s13098-023-01096-9. Q2.

Insulin Requirement for Gestational Diabetes Control Is Related to Higher Vitamin D Levels up to 1 Year Postpartum: A Prospective Cohort. Molina-Vega M, Picón-César MJ, Lima-Rubio F, Gutiérrez-Repiso C, **Linares-Pineda TM**, Suárez-Arana M, Fernández-Ramos AM, Tinahones FJ, Morcillo S. *Antioxidants* · Nov 11, 2022. *Antioxidants* · Nov 11, 2022. I.F: 7.675. D1.

Epigenetic changes in the metabolically healthy obese: A case-control versus a prospective study. **Linares-Pineda TM**, Boughanem H, Gutiérrez-Repiso C, Macías-González M, Andrés-León E, Rojo-Martínez G, Valdés S, Tinahones FJ, Morcillo S. *European Journal of Clinical Investigation* · Aug 1, 2022. *European Journal of Clinical Investigation* · Aug 1, 2022. IF: 5.722. Q1.

Impact of the Gestational Diabetes Diagnostic Criteria during the Pandemic: An Observational study. Molina-Vega M, Gutiérrez-Repiso C, Lima-Rubio F, Suárez-Arana M, **Linares-Pineda TM**, Cobos Díaz A, Tinahones FJ, Morcillo S, Picón-César

MJ. Journal of Clinical Medicine · Oct 24, 2021 Journal of Clinical Medicine · Oct 24, 2021. IF: 4.964. Q2.

Epigenetic Biomarkers of Transition from Metabolically Healthy Obesity to Metabolically Unhealthy Obesity Phenotype: A Prospective Study. Gutiérrez-Repiso C, **Linares-Pineda TM***, Gonzalez-Jimenez A, Aguilar-Lineros F, Valdés S, Soriguer F, Rojo-Martínez G, Tinahones FJ, Morcillo S. International Journal of Molecular Science · Sep 27, 2021 International Journal of Molecular Science · Sep 27, 2021. *Co-author. IF: 6.208. Q1.

“A ti que te sentí durante unos meses,
A ti que, aunque no estes, estuviste,
A ti que nunca te olvidaremos,
A ti, Félix”

*“Omnia vincit amor et nos cedamos amori”
Bucólicas, Virgilio*

Agradecimientos

Este apartado no ha sido fácil de escribir. En estos últimos cuatro años han pasado muchas cosas, incluida una pandemia mundial que coincidió de lleno con el inicio de mi etapa predoctoral. Este tiempo ha estado cargado de sentimientos. Muchos ya los conocía de antes, otros han sido totalmente nuevos. Pandemias, cuarentenas, funerales, bodas, viajes, nacimientos, hospitalizaciones, todo eso en paralelo a llevar a cabo esta tesis doctoral. No, no es fácil escribir estas líneas rememorando a todas esas personas a las que debo agradecer que a día de hoy pueda presentar este trabajo.

Pero empecemos por el principio. En primer lugar, tengo que dar las gracias a mis directores de tesis que han hecho posible la existencia de este trabajo: la Dra. Sonsoles Morcillo, el Dr. Antonio Sánchez Pozo y la Dra. María José Picón. Al Dr. Sánchez Pozo, gracias por haber querido participar en este proyecto. Por guiarnos a nivel académico, y haber sido pieza clave en poder hacer la tesis en la Universidad de Granada. Agradecer a la Dra. Picón, ya que sin ella este proyecto no sería el mismo, por no decir que no sería. Le debemos el reclutamiento de las pacientes y ha sido clave en la idea del estudio. Gracias por estar cuando necesitábamos apoyo clínico y nos surgían dudas. En definitiva, te doy las gracias por aportar tu visión clínica a esta porción del proyecto más molecular, potenciando la transversalidad de la línea de investigación.

Gracias a la Dra. Sonsoles Morcillo por elegir mi curriculum entre los muchos que seguro que le llegaron. Por haberme dado la oportunidad de pedir la beca PFIS sin conocerme de antes. Y sobre todo quiero darte las gracias por ser algo más que una simple directora de tesis, has sido

mentora y confidente, has tenido mucha paciencia con mi posible dislexia no diagnosticada y me has entendido y comprendido cada vez que he tenido algún problema personal. Me has enseñado a hacer ciencia, a ser resolutiva y a encontrar respuestas a todas tus preguntas (que no son pocas). Me has dado ese empujón cuando lo necesitaba y a ver que en la vida se tienen que tener prioridades. Además, por mi experiencia puedo decir que eres una gran científica y un ejemplo de liderazgo, pero sobre todo eres una gran persona. No he podido tener más suerte cayendo en tu grupo y en concreto en tu línea de investigación.

No podría seguir sin mencionar a todas las personas que han estado codo a codo en el laboratorio. El grupo de endocrino no es un grupo pequeño, y a lo largo de estos cuatro años han ido cambiando algunas caras. Pero como hice anteriormente, empezaré por el principio. Empezaré por aquellas personas que me enseñaron las primeras técnicas en el laboratorio, y a moverme por él con soltura. Por un lado, Ali, cuánto se la echa de menos por el laboratorio. Me enseñaste cómo funcionaba todo por el labo en esas primeras semanas. También agradecer a Francisca que me enseñara el procesamiento de las muestras y a manejarme con los protocolos. No puedo olvidarme de la Dra. Carolina Gutiérrez-Repiso, “Caro” para todos. En los primeros meses fue la que me enseñó muchos protocolos de biología molecular y tuvo la paciencia para ver de dónde podían salir las malditas contaminaciones. Entre ella y Sonsoles, me enseñaron todo lo que sé hoy en día de qPCR y de la interpretación de cada curva que nos daba el aparato. En este párrafo de inicios, también me gustaría agradecer al Dr. Eduardo Andrés-León. Tutor de mi TFM en bioinformática, me enseñó las herramientas básicas para hacer los análisis de metilación. Tengo que

agradecerle que me supo enseñar y guiar para poder hacer los distintos análisis bioinformáticos que soy capaz de hacer hoy por hoy. Tanto Caro como Eduardo sentaron las bases de todo lo que he ido desarrollando en esta etapa.

Seguir dando las gracias a Virginia, Elena, Alba S. y Ana porque siempre han estado ahí cuando las he necesitado (tanto nivel científico como personal). Por todos esos ratos de charlas y de desahogo. Sin vosotras esta tesis no hubiera sido la misma, sois más que compañeras de trabajo. No me olvido de los juveniles, Gema, Jesús, Carlos, Josema, Belén, María José, Alba E., Raquel y a Nerea. Trajeron aire fresco al laboratorio y hace que sea muy fácil trabajar con cada uno de ellos. Les agradezco todos los ratos fuera del lab que dan la vida. A Nerea le tengo que hacer una mención especial, cuando llegaste te enseñé lo básico y me fui para no volver en casi un mes por el maldito accidente. No puedo agradecerte más que hayas estado ahí cubriéndome las espaldas cada vez que me pasaba algo. Y qué decir de lo bien que nos lo hemos pasado este último año yendo a congresos, Valencia, Hamburgo y Sevilla. Nos echaremos de menos cuando te vayas a Madrid.

No me olvido de los demás compañeros del grupo, a los doctores: Patri, Isa, Manuel, Violeta, Almudena, Hatim, Mora, Libia y Lourdes. Entre todos, me habéis enseñado cómo funciona la ciencia, a que a veces es frustrante pero luego las cosas van saliendo. Una mención especial al jefe Dr. Francisco J. Tinahones, Paco para todos en el grupo. Agradecer haber podido ser parte de un grupo tan grande y multidisciplinar, agradecerle también su apoyo, sus ganas de que salgan todos los proyectos adelante y que sigamos

creciendo. Nunca olvidaré todas las convivencias que hacen que nos entendamos un poco mejor como equipo.

Agradecer a la Dra. Christine Sommer, por permitirme ir a su grupo de investigación a realizar la estancia de tres meses. También agradecerle su implicación a la hora de enseñarme nuevas herramientas informáticas, su cercanía y bienvenida al grupo. A los doctores Kåre Inge Birkeland y Elisabeth Qvigstad por enseñarme su punto crítico y clínico a la hora de plantear los análisis de los datos. Entre todos me enseñaron que la cultura noruega no es tan fría como los demás pensamos. No me olvido de Nicolás, y él seguro que no se olvidará de mí por pesada. Tres meses que iba cada vez que tenía un problema para que me ayudara, a que me explicara las distintas cosas que iba necesitando. Desde luego gran parte de lo que aprendí en esos meses en Oslo se lo debo a él. En este párrafo sobre Oslo, no puedo olvidarme de los españoles que conocí Paz, Álvaro, Alicia... Entre todos hicisteis que la estancia fuera mucho más llevadera.

Y hasta aquí he dado las gracias a todos los que me han ayudado en la parte más profesional de estos cuatro años. Pero esto no es todo, tengo que agradecer al apoyo incondicional de mis amigos y familiares. Tengo que agradecer a Irene, mi compi de piso. Hemos compartido tantos ratos, tantas charlas, películas, series... en definitiva vida juntas. Hemos ido casi a la par en el camino de la tesis, tú con la tuya con tus idas y venidas yo con las mías propias. Que decirte que no sepas, que eres un gran apoyo en mi día a día aquí en Málaga, y he tenido mucha suerte de tener una compi tan genial, empática y cariñosa. Millones de gracias por cuidar de Akira (aunque seas alérgica a los gatos), por acompañarme en la elección de vestido de novia y por estar en los momentos que necesitaba. Eres de esas

personas que sabes que si las llamas en el último momento no te van a fallar. Gracias por estar siempre ahí.

Ni que decir de mis amigas de toda la vida: Inma, Ana, Helena, Marta, Cristina y Tere. Una parte muy importante en mi vida, sin ellas mi salud mental no sería la misma. Les tengo que agradecer su apoyo en todo momento, y por estar ahí siempre que lo he necesitado. Aunque a veces la distancia haya hecho más complicado mantener el contacto, sé que cuando os vuelvo a ver es como si no hubiera pasado el tiempo. Todas sois ejemplo de grandes personas, habéis conseguido todo aquello que os propusisteis, sois capaces de todo y sois unas amigas increíbles. No puedo tener más suerte de teneros.

Y poco a poco vamos llegando al final de los agradecimientos. No sin antes agradecer a mi familia, empezando por mis tías y primos y terminando por mis abuelos, mis padres y mi hermana. Este apartado es un pelín más triste ya que mi abuelo le hubiera encantado estar en este momento, pero se fue cuando tenía que irse. Sin mis abuelos esto no se habría escrito por su apoyo incondicional desde el momento, y aunque mi abuela no termine de entender lo que hago, sabe que es importante y se interesa por lo que hago. Mis padres, amor incondicional y apoyo continuo. Cuando todo falla se que puedo contar con ellos. Y mi hermana, relación a distancia todos nuestro veintitantos y ahora casi treinta y tantos, pero siempre sé que puedo contar contigo cuando tengo cualquier problema, no se que hubiera sido de mi cuando pasó lo de Costa Rica. Sabes con quien hablar cuando pasan cosas malas, y gracias a ti pudo venir Silvia para que no estuviera sola. No solo te estoy agradecida por esto, estoy agradecida porque me has enseñado mucho como hermana mayor, y mucho de lo que soy te lo debo a ti. Bueno

y terminar esta parte familiar dando las gracias a mi prima Silvia, gracias por haber venido cuando pasó todo, ese martes de incertidumbre de no saber que iba a pasar, fuiste de gran apoyo y compañía la representante de toda la familia que no pudo viajar. Te debo un viaje a EE. UU..

Bueno y no puedo terminar este apartado sin darle las gracias a Jesús. Compañero de vida, amigo, cómplice y confidente. En estos casi diez años hemos vivido muchas cosas juntos, has estado ahí en momentos clave de mi vida y siempre he sentido tu amor y apoyo. Sé que siempre puedo contar contigo cuando mi mente es un caos, y me ayudas a poner en orden mi vida. Eres un ejemplo de superación y resiliencia. Esta tesis no se entiende sin el amor que sentimos el uno por el otro. Sé que aún nos queda mucho por construir, pero por todo lo que tenemos ahora gracias.

En resumen, gracias a todos los que me habéis acompañado en este camino. Espero que muchos me sigáis acompañando el resto de mi vida.



Solen, Artista: Munch. Aula magna de la Universidad de Oslo, Oslo (Noruega).

Tabla de contenido

Abstract	22
Resumen	26
Introducción	32
Un problema de salud en crecimiento	33
Diabetes gestacional.....	34
Epidemiología	35
Criterios diagnósticos	36
Mecanismos moleculares.....	37
Consecuencias	41
Factores de riesgo.....	43
Factores genéticos.....	45
Epigenética	45
Mecanismos epigenéticos	46
Tecnología para la detección de metilación del ADN	51
Técnicas basadas en la conversión con bisulfito.....	51
Procesamiento y finalidad de los estudios EWAS.....	55
Flujo de trabajo en el análisis bioinformático de los EWAS.....	56
Obtención y carga de los datos	56
Control de calidad.....	57
Normalización.....	57
Filtrado de las sondas	58
Eliminación de la variación debida a la técnica	58
Tipos de análisis de EWAS	59
Obtención de las DMPs	59
Obtención de las DMRs	60
Estudios de enriquecimiento funcional.....	61
Cálculo de puntuación del riesgo de metilación o methylation risk score ...	63

Análisis de loci de rasgos cuantitativo de metilación (m-QTLs)	64
Hipótesis y objetivos	66
Hipótesis.....	67
Objetivos.....	67
Materiales y métodos	70
Contexto	71
Sujetos del estudio	71
Procedimiento y variables cohorte EPIDG.....	72
Cohorte EPIPREG	76
Procesamiento de las muestras y determinación de las variables de estudio	77
Procesamiento de muestras:.....	77
Determinaciones bioquímicas	78
Determinación de variables clínicas y antropométricas.....	78
Aislamiento de ADN y bisulfitación	79
Análisis de la metilación del ADN.....	79
Análisis estadístico.....	81
Estadística descriptiva de las cohortes	81
Obtención β /M-values y extracción de los DMPs.....	81
Análisis de los datos obtenidos del array de metilación en sangre materna en EPIDG y EPIPREG.....	82
Análisis de los datos obtenidos del array de metilación en saliva de la descendencia EPIDG	82
Análisis funcional.....	84
mQTLs.....	85
Resultados	86
Objetivo 1- Estudio de los patrones de metilación del ADN en muestras de sangre periférica de mujeres gestantes con y sin DG, tanto en el momento de diagnóstico de DG como al final del embarazo	88
1.1-..... Identificación de marcas epigenéticas involucradas en rutas relacionadas con DG.....	90
1.2- Marcas epigenéticas asociadas con la DG en dos momentos durante el embarazo.....	101

Objetivo 2- Comparación del perfil epigenético en muestras de saliva entre la descendencia nacida de madres con DG y la de madres sin DG.	116
Características de la cohorte	118
Análisis de metilación diferencial	120
Análisis enriquecimiento funcional	122
Objetivo 3- Evaluación de la relación de un MRS de DM2 con la DG	126
Discusión	136
El patrón de metilación del ADN en sangre periférica de mujeres gestantes con y sin DG es diferente y afecta a numerosos genes asociados a la DMT2	137
Existe un patrón epigenético común de metilación a lo largo del embarazo que diferencia entre DG y no-DG.	141
Potenciales biomarcadores epigenéticos de DG	143
Asociación de un MRS de DMT2 con la DG	148
Comentario final	152
Conclusiones	155
Conclusions	157
Bibliografía.....	160
Material suplementario	188
Tablas suplementarias:	189
Figuras suplementarias:.....	206
Abreviaturas.....	215

Abstract

The increasing prevalence of obesity and type 2 diabetes in recent years has become one of the major global health problems. Obesity raises the risk of developing various diseases, including type 2 diabetes (T2DM), hypertension and cardiovascular diseases, among others. Type 2 diabetes is characterized by insulin deficiency caused by beta cell dysfunction and insulin resistance in target organs. Despite the advances in medical sciences and prevention and health promotion, the prevalence of these conditions and their comorbidities continues to rise.

Women with gestational diabetes (GDM) have an increased risk of developing T2DM. Moreover, T2DM is known to increase the risk of developing obesity in the offspring. Therefore, gestational diabetes can help us understand how an adverse environment, such as hyperglycaemia environment, can affect to the metabolism of both mothers and offspring. It is believed that DNA methylation is influenced by factors such as diet and the intrauterine environment, which may impact on the development of gestational diabetes. Thus, throughout this work, we have studied how epigenetic marks may influence on the development of gestational diabetes, type 2 diabetes, and other metabolism disorders.

To achieve this, we have divided this thesis into three main objectives:

The **first objective** was to identify alterations in epigenetic profile in women with GDM. Within this objective, we sought epigenetic marks that may be related to the development of gestational diabetes. For this purpose, we selected a total of 32 pregnant women from the EPIDG cohort, 16 having gestational diabetes and the other 16 without gestational diabetes. We obtained the DNA methylation pattern from the Illumina

Methylation Epic BeadChip, using peripheral blood samples from the diagnostic visit (weeks 26-28). Differentially methylated positions (DMPs) were extracted using the ChAMP and limma packages in R 2.9.10, with a threshold of $FDR < 0.05$, $\Delta\beta > |5\%$ and $B > 0$. We found a total of 1,141 DMPs, of which 714 were annotated in genes. A functional analysis was conducted, and 23 genes significantly related to carbohydrate metabolism were identified. Finally, a total of 27 DMPs were correlated with biochemical variables such as glucose levels at different points of the oral glucose tolerance test, fasting glucose, cholesterol, HOMA-IR, and HbA1c at different visits during pregnancy and postpartum.

On the other hand, we identified epigenetic marks that remained stable throughout pregnancy, which could be of interest as potential biomarkers. For this, we used the following methodology: we conducted a whole epigenome-wide association study (EWAS) in maternal peripheral blood samples from 32 pregnant women (16 with GDM and 16 without GDM) at pregnancy week 24-28 and 36-38. We collected biochemical, anthropometric, and obstetrical variables from all participants. The main results were validated in an independent cohort EPIREG with different ethnic backgrounds (Europeans = 307; South Asians = 165). A total of 272 CpG sites remained significantly different between pregnant women with GDM and those without GDM at two time points during pregnancy. The significant CpG sites were related to pathways associated with type 1 diabetes, insulin resistance, and insulin secretion. Cg01459453 (SELP gene) was the most differentiated in the GDM group compared to the non-GDM group (73.6 vs. 60.9, $p = 1.06E-11$; $FDR = 7.87E-06$). Three CpG sites (cg01459453, cg15329406, and cg04095097) could discriminate between

GDM cases and non-GDM (AUC = 1; $p = 1.26E-09$). Three differentially methylated positions were replicated in an independent cohort.

The **second objective** was to analyse the epigenetic profile in offspring during the first year of life and study if there was a differential methylated pattern between offspring born to mothers with gestational diabetes compared to offspring of controls. For this objective, the methodology used was as follows: we conducted an EWAS from saliva samples in a total of 28 children from the EPIDG cohort, 15 of whom were born to mothers without GDM and 13 to mothers with GDM. Saliva samples were collected two weeks after birth (T2) and one year later (T3). Differentially methylated positions (DMP) were obtained using the limma package for R Studio, adjusted for newborn sex, GDM treatment (diet or insulin), and maternal weight gain ($P_{val} < 0.01$). A total of 5,132 DMPs were found in the postpartum period, and 6,960 DMPs at one year after birth. Sixteen of these DMPs were common at both time points (T2 and T3) in the same direction. At T2, the methylation levels of 4 CpGs correlated significantly with weight percentile, head circumference, and height. Finally, at T3 the methylation values of 6 CpGs were also related to those percentiles.

Finally, the **third objective** aimed to assess a Methylation Risk Score (MRS) for type 2 diabetes and the risk of developing GDM in different populations and using different diagnostic criteria. The methodology was as follows: In two independent pregnancy cohorts, EPIPREG ($n=480$) and EPIDG ($n=32$), DNA methylation in peripheral blood leukocytes was measured at a gestational age of 28 \pm 2. We constructed an MRS in EPIPREG and EPIDG based on CpG hits from a meta-analysis published of EWAS of T2DM. With mixed model logistic regression of EPIPREG and EPIDG, MRS for T2DM was

associated with GDM risk: odds ratio (OR) [95% CI]: 1.3 [1.1-1.8], P=0.002 for the unadjusted model, and 1.4 [1.1-1.7], P=0.00014 for the adjusted model. Furthermore, we found 6 CpGs (cg14020176, cg22650271, cg14870271, cg27243685, cg06378491, cg25130381) associated with GDM, and some of their methylation quantitative loci (mQTLs) were related to T2DM and GDM.

In conclusion, our findings suggest that there is a differentiated methylation pattern between GDM and non-GDM. Additionally, the genes annotated to the DMPs may be involved in the development of GDM and alterations in related metabolic variables. Moreover, there are epigenetic marks that differed between GDM cases and controls during pregnancy, which could potentially serve as biomarkers for the diagnosis or prediction of GDM. We have also identified a distinct methylation pattern in saliva samples between children born to mothers who experienced GDM during pregnancy compared to children born to mothers without GDM. Some of these DMPs persist at the age of one year and are associated with variables related to the child's growth pattern. Finally, we demonstrate that DNA methylation marks for T2DM are associated with GDM, suggesting shared epigenetic mechanisms between GDM and T2DM.

Resumen

La creciente prevalencia de la obesidad y la diabetes tipo 2 en los últimos años se ha convertido en uno de los principales problemas de salud a nivel mundial. La obesidad aumenta el riesgo de desarrollar diversas enfermedades, incluyendo la diabetes tipo 2, hipertensión, enfermedades cardiovasculares entre otras. La diabetes tipo 2 se caracteriza por una deficiencia de insulina causada por la disfunción de la célula beta y la resistencia a la insulina en los órganos diana. A pesar de los avances médicos y las medidas de prevención, la prevalencia de estas condiciones y sus comorbilidades sigue en aumento.

En este contexto, la Diabetes Gestacional (DG) nos puede ayudar a entender cómo un ambiente adverso puede afectar a la salud tanto de la madre como de la descendencia, ya que se ha visto que tienen mayor probabilidad de desarrollar diabetes tipo 2, obesidad y otras complicaciones metabólicas.

Se sabe que la metilación del ADN está influenciada por factores ambientales como pueden ser la nutrición y el entorno intrauterino, lo que puede estar influyendo en el desarrollo de la diabetes gestacional. Por ello a lo largo de este trabajo hemos estudiado cómo el ambiente adverso ocasionado por la hiperglucemia puede afectar al epigenoma y a su vez, estas marcas epigenéticas pueden estar influyendo en el desarrollo de la diabetes gestacional, diabetes tipo 2 u otras alteraciones del metabolismo de la glucosa.

Para ello, hemos dividido esta tesis en **3 objetivos** principales:

El **primer objetivo** era conocer el perfil epigenético diferencial entre madres con diabetes gestacional y controles. Dentro de este objetivo, por un lado, hemos buscado marcas epigenéticas que puedan estar relacionadas con el desarrollo de la diabetes gestacional. Para este objetivo hemos seguido la siguiente metodología: se seleccionaron un total de 32 mujeres embarazadas de la cohorte EPIDG, 16 con diabetes gestacional y las otras 16 sin diabetes gestacional. Se obtuvo el patrón de metilación del ADN a partir del Illumina Methylation Epic BeadChip, utilizando muestras de sangre periférica de la visita diagnóstico (semana 26-28). Se extrajeron posiciones diferencialmente metiladas utilizando los paquetes ChAMP y limma en R 2.9.10, con un umbral de FDR <0.05 , $\Delta\beta > |5|\%$ y $B > 0$. Se encontraron un total de 1.141 DMPs, de las cuales 714 estaban anotadas en genes. Se realizó un análisis funcional y se identificaron 23 genes relacionados significativamente con el metabolismo de los carbohidratos. Finalmente, se encontraron un total de 27 DMPs correlacionados con variables bioquímicas como los niveles de glucosa en diferentes puntos de la prueba de tolerancia a la glucosa oral, colesterol, HOMAIR y HbA1c en diferentes visitas durante el embarazo y el parto.

Por otro lado, hemos identificados marcas epigenéticas que se mantienen estables a lo largo del embarazo pudiendo ser de interés como potencial biomarcador. Para ello, hemos utilizado la siguiente metodología: Realizamos un estudio de asociación de todo del epigenoma (EWAS) en muestras de sangre periférica materna de 32 mujeres embarazadas (16 con DG y 16 sin DG) en las semanas de embarazo 24-28 y 36-38. Se recopilaron variables bioquímicas, antropométricas y obstétricas de todas las participantes. Los principales resultados se validaron en una cohorte

independiente con diferente origen étnico (europeas = 307; asiáticas del sur = 165). Doscientos setenta y dos sitios CpG permanecieron significativamente diferentes entre las mujeres embarazadas con DG y sin DG en dos momentos durante el embarazo. Los sitios CpG significativos estaban relacionados con rutas metabólicas asociadas con la diabetes mellitus tipo I, resistencia y secreción de insulina. El sitio cg01459453, anotado al gen SELP, fue el más diferenciado en el grupo con DG en comparación con el grupo sin DG (73.6 vs. 60.9, $p = 1.06E-11$; FDR = 7.87E-06). Tres sitios CpG (cg01459453, cg15329406 y cg04095097) pudieron discriminar entre casos de DG y controles (AUC = 1; $p = 1.26E-09$). Tres posiciones con metilación diferencial se replicaron en una cohorte independiente.

El **segundo objetivo** era conocer el perfil epigenético en la descendencia a lo largo del primer año de vida y ver si había un patrón diferencial entre la descendencia de nacidos de madres con diabetes gestacional frente a la descendencia de los controles. Para este objetivo la metodología utilizada ha sido la siguiente: se realizó un EWAS a partir de muestras de saliva en un total de 28 niños de la cohorte EPIDG, de los cuales 15 nacieron de mujeres sin DG y 13 de mujeres con DG. Se tomaron muestras de saliva dos semanas después del nacimiento (T.2) y un año después (T.3). Se obtuvieron posiciones con metilación diferencial (DMP) utilizando el paquete limma para R Studio, ajustadas por el sexo del recién nacido, el tratamiento de DG (dieta o insulina) y el aumento de peso materno ($P.val < 0.01$). Se encontraron un total de 5.132 DMPs en el período postparto y 6.960 DMPs un año después del nacimiento. Dieciséis de estos DMPs fueron comunes en ambos puntos (T2 y T3) en la misma dirección.

Un análisis de correlación con estos DMPs mostró que los valores de metilación en T2 de 4 CpGs estaban correlacionados con los percentiles del percentil de peso, la circunferencia de la cabeza y la altura. Los valores de metilación en T3 de 8 CpGs también estaban relacionados con esos percentiles. Finalmente, 4 CpGs en T2 estaban relacionados con el peso en T3.

Por último, en el **tercer objetivo** se quiso evaluar si un *Methylation Risk Score* (MRS) de diabetes tipo 2 (DM2) se asociaba también con la DG, en distintas poblaciones y con criterios diagnósticos distintos. La metodología a seguir fue la siguiente: En dos cohortes de embarazo independientes, EPIPREG (n=480) y EPIDG (n=32), se midió la metilación del ADN en leucocitos de sangre periférica a una edad gestacional de 28±2 semanas. MRS en EPIPREG y EPIDG basado en los sitios CpG de un estudio de asociación de todo el epigenoma (EWAS) publicado sobre la diabetes tipo 2. Con modelos de regresión logística mixta, el MRS para diabetes tipo 2 se asoció con la probabilidad de tener DG, en ambas cohortes, con un odds ratio (OR) [IC del 95%]= 1.3 [1.1-1.8], en el modelo sin ajustar, y OR=1.4 [1.1-1.7], cuando ajustamos por edad, IMC pre-embarazo, historia familiar de diabetes y fumadora. Además, encontramos 6 sitios CpG (cg14020176, cg22650271, cg14870271, cg27243685, cg06378491, cg25130381) asociados con DG, y algunos de sus loci cuantitativos de metilación (mQTLs) estaban relacionados con DM2 y DG.

En conclusión, nuestros hallazgos sugieren que existe un patrón de metilación diferenciado entre DG y no DG. Los genes anotados en los DMPs pueden estar involucrados en el desarrollo de la DG y en alteraciones de variables metabólicas relacionadas. Además, existen marcas epigenéticas

que difirieron entre los casos de DG y los controles durante el embarazo, lo que podría servir potencialmente como biomarcadores para la predicción de la DG, si se demostrara en etapas más tempranas del embarazo. También hemos identificado un patrón distintivo en las muestras de saliva entre los niños nacidos de madres que experimentaron DG durante el embarazo en comparación con los niños nacidos de madres sin DG. Algunos de estos DMPs persisten a la edad de un año y están asociados con variables relacionadas con el patrón de crecimiento del niño.

Finalmente, demostramos que algunas marcas de metilación del ADN para la diabetes tipo 2 están asociadas también con la DG, lo que sugiere mecanismos epigenéticos compartidos entre la DG y la DM2.

Introducción

Un problema de salud en crecimiento

En los últimos años, la creciente prevalencia de la obesidad y la diabetes mellitus tipo 2 (DM2) se han convertido en uno de los principales problemas de salud en el mundo. Se estima que más de mil millones de personas en todo el mundo sufren de obesidad. En Europa, casi el 60% de los adultos y cerca del 29% de los niños y el 27% de las niñas padecen de sobrepeso u obesidad (1). La obesidad aumenta el riesgo de desarrollar otras enfermedades como la DM2, hipertensión, enfermedades cardiovasculares, musculoesqueléticas, renales crónicas y enfermedades mentales, entre otras (2). La DM2 se caracteriza por una deficiencia relativa de insulina causada por la disfunción de las células beta y la resistencia a la insulina en los órganos diana (3). Se estima que el 8.3% de la población mundial de 20 a 79 años tiene DM2, y se prevé que aumentará al 10.2% para el año 2030 (4). A pesar del avance en el conocimiento de los mecanismos implicados, y la implementación de medidas de prevención para frenar su progresión, la realidad es que esta condición y sus comorbilidades continúan en aumento.

La obesidad y la diabetes adquieren una especial relevancia clínica, ya que los hijos de madres con obesidad y/o diabetes tienen más probabilidades de desarrollar estas patologías u otras enfermedades metabólicas asociadas a lo largo de sus vidas. Las razones no están claras. Algunas teorías, como la conocida hipótesis DOHAD (de sus siglas en inglés Developmental Origins of Health and Disease) y la reciente teoría de los primeros 1000 días, propone que la exposición a un entorno adverso durante el período periconcepcional y/o intrauterino confiere un mayor riesgo de por vida para el desarrollo de enfermedades metabólicas o

trastornos relacionados en la descendencia. Durante los primeros mil días de vida, periodo que abarca desde la concepción hasta los dos primeros años de vida de una persona, existe una capacidad adaptativa a diferentes estímulos y entornos sociales que pueden ser determinantes para el futuro del individuo (5,6).

La diabetes mellitus gestacional (DG) es un marco ideal para estudiar posibles biomarcadores para el desarrollo de futuros trastornos metabólicos tanto en mujeres embarazadas como en su descendencia, así como los mecanismos implicados. Varios estudios han demostrado un mayor riesgo de desarrollar DM2 en mujeres embarazadas con DG en comparación con mujeres embarazadas sin diabetes (7). En uno de ellos, realizado en indios Pima demostró, que los hermanos nacidos después de que la madre desarrollara diabetes tenían un mayor riesgo de DM2 en comparación con aquellos nacidos antes de que la madre desarrollara diabetes, destacando la importancia del entorno intrauterino diabético como un factor de mayor riesgo que la susceptibilidad genética (8). Por lo tanto, aunque la DG es una condición transitoria, se considera un factor de riesgo para la salud tanto del recién nacido como de la madre.

Diabetes gestacional

La DG queda conformada por todas aquellas situaciones de hiperglucemia no pre-existentes (o conocidas) previas a la gestación pero que se pueden presentar de forma diversa (9,10). Es la complicación más común durante el embarazo y se asocia con acontecimientos adversos a corto plazo tales como recién nacido prematuro, cesárea o preeclampsia, y a largo plazo DM2, síndrome metabólico, y enfermedades cardiovasculares, entre otros. Además, la descendencia de madres con DG tiene mayor riesgo de

presentar macrosomía en el momento del nacimiento, mayor masa grasa durante la infancia y, en la edad adulta, mayor riesgo de obesidad (11).

Epidemiología

A nivel mundial es complicado comparar la prevalencia de la DG ya que existe gran variedad de estrategias de detección y de criterios diagnósticos utilizados para identificar a las mujeres con DG. La prevalencia puede variar entre el 1 y 30% dependiendo de la población de estudio y de los criterios diagnósticos utilizados (12).

Global prevalence of Gestational diabetes



Figura 1- Prevalencia global de DG.

Según la Federación Internacional de Diabetes (IDF de sus siglas en inglés), la prevalencia estandarizada a nivel mundial ronda el 14%. Las tasas de prevalencia por región fueron las siguientes: 7.1% en América del Norte y el Caribe, 7.8% en Europa, 10.4% en América del Sur y América Central,

14.2% en África, 14.7% en el Pacífico Occidental, 20.8% en Asia Sudoriental y 27.6% en Oriente Medio y el Norte de África (Figura 1) (13).

En España, la prevalencia se estima en torno al 8,8% utilizando los criterios del Grupo Nacional de Datos de la Diabetes (National Diabetes Data Group, NDDG del inglés). La prevalencia aumentaría en torno a un 3% si se utilizaran otros criterios diagnósticos como los de Carpenter y Coustan (CC) o de la Asociación Internacional del Grupo de Estudio sobre Diabetes y Embarazo (del inglés, IADPSG). Sin embargo, el Grupo Español de Diabetes y Embarazo (GEDE) acordó seguir aplicando los criterios del NDDG porque el objetivo del diagnóstico y tratamiento de la DG es la prevención de la morbilidad perinatal. Aplicando criterios más estrictos como los de la IADPSG, se identificaría un grupo de gestantes con DG, pero con baja morbilidad perinatal. Según el GEDE, tal aumento de la prevalencia no sería aceptable para el sistema sanitario público (14).

Criterios diagnósticos

El principal problema para el diagnóstico de la DG es la falta de consenso. En la tabla 1 se muestran los distintos criterios diagnósticos.

NDDG y CC establecen, para el diagnóstico de la DG en las semanas 24-28 de gestación, dos pasos: un primer paso de screening haciendo un test de tolerancia a la glucosa de 50gr (test de O'Sullivan). En el caso de un resultado positivo (glucosa después de 1 hora por encima de 7.8 mmol/L o lo que es lo mismo 140 mg/dl), se lleva a cabo una sobrecarga oral de glucosa (SOG) de 100g y se recogen muestras de sangre en ayunas, a la hora, dos horas y a las 3 horas. En la tabla 1 se recogen los criterios más utilizados para el diagnóstico (15,16) (Tabla 1), pero estos no son los únicos.

Tabla 1- Diferentes criterios diagnósticos aceptados.

Criterio	Procedimiento	Puntos de corte de los niveles de glucosa mmol/L (mg/dL)				N
		Ayuno	1h	2h	3h	
NDDG	O'Sullivan test (SOG 50g)/SOG 100g, 3h	5.8 (104)	10.6 (190)	9.2 (165)	8.1 (145)	2 o más valores anormales
CC	O'Sullivan test (SOG 50g)/SOG 100g, 3h	5.3 (95)	10 (180)	8.6 (154)	7.8 (140)	2 o más valores anormales
IADPSG	No O'Sullivan/SOG 75g, 2h	5.1 (91)	10 (180)	8.5 (153)		Solo 1 valor

N: número de puntos positivos por encima del umbral. NDDG: National Diabetes Data Group. CC: Carpenter y Coustan. IADPSG: the International Association of the Diabetes and Pregnancy Study Groups.

En 2010, la IADPSG desarrolló un nuevo criterio diagnóstico basado en los resultados del estudio Hiperglucemia y efectos adversos durante el embarazo (HAPO, del inglés). Se trata de un estudio multicéntrico llevado a cabo con más de 25.000 mujeres gestantes a las que se le hizo un seguimiento para comprobar la relación entre la glucemia materna y la adiposidad fetal. Llegaron a la conclusión de que existe una correlación lineal entre ambos parámetros y que no se podría establecer un punto de corte. Por ello, la IADPSG propuso coger un punto de corte arbitrario en la curva de glucosa que corresponde a la OR de 1.75 para los principales *end-point*. Este nuevo criterio se basa en un screening de un solo paso, con una sobrecarga oral de 75gr en las semanas 24-28 de embarazo. La Asociación de Diabetes Americana (ADA) en 2011 y la Organización Mundial de la Salud (OMS) a partir de 2013 empezaron a seguir este criterio (17,18).

Mecanismos moleculares

La DG se caracteriza principalmente por una resistencia a la insulina y defectos en la función de las células β .

En un **embarazo normal**, la producción de glucosa aumenta un 30% para cubrir las necesidades energéticas. Sin embargo, el nivel de glucosa en sangre disminuye en ayunas en el torrente sanguíneo. Esto se debe muy probablemente a dos factores principales: un aumento del volumen plasmático al principio del embarazo y un aumento posterior de la utilización de glucosa por la unidad fetoplacentaria a medida que avanza el embarazo. Más adelante en la gestación, la sensibilidad periférica a la insulina experimenta un descenso significativo de aproximadamente el 50%, esta resistencia probablemente se produzca para incrementar el aporte de glucosa al feto. La sensibilidad periférica a la insulina se refiere a la capacidad de la insulina para potenciar la captación de glucosa en los músculos esqueléticos y el tejido adiposo (19). En condiciones normales, la reducción de la sensibilidad a la insulina desencadena una respuesta compensatoria en forma de un aumento 2-3 más secreción de insulina. Este aumento de la secreción de insulina ayuda a mantener estables los niveles de glucemia. Esta alteración de la sensibilidad a la insulina durante el embarazo afecta, además, al metabolismo de los aminoácidos y los lípidos maternos (20). La resistencia a la insulina durante el embarazo se ha asociado al efecto de las hormonas liberadas por la placenta (HPL), como el lactógeno placentario y la hormona de crecimiento placentario (PGH). En ocasiones, factores de riesgo de DG como la obesidad y otros factores inflamatorios pueden alterar la cascada de señalización de la insulina. Se sabe que el factor de necrosis tumoral (TNF) desencadena una vía de señalización que aumenta la producción de esfingomielinasa y ceramida, lo que dificulta la autofosforilación en tirosina del receptor de insulina y facilita la fosforilación en serina del sustrato 1 del receptor de insulina (IRS1), que altera la cascada de señalización de la insulina. Durante

el embarazo, existe una relación inversa significativa entre los niveles circulantes de TNF y la sensibilidad a la insulina, incluso teniendo en cuenta la masa grasa materna (21–23).

En **mujeres no embarazadas con tolerancia normal a la glucosa**, cuando la insulina se une al receptor de insulina en la superficie de las células de tejidos periféricos como el músculo esquelético, desencadena la captación de glucosa por las células. Esta interacción activa la autofosforilación, a través del dominio tirosina quinasa, de la subunidad β del receptor de insulina (IR β), iniciando una cascada de señalización. Esta cascada conduce a la traslocación de los transportadores de glucosa tipo 4 (GLUT4) a la superficie celular, facilitando la captación de glucosa por la célula (24). La resistencia a la insulina surge cuando las células dejan de responder adecuadamente a los efectos de la insulina. Este fallo en la señalización de la insulina se produce a nivel molecular y conduce a una translocación insuficiente de GLUT4 a la membrana plasmática. Durante el embarazo, la sensibilidad a la insulina disminuye a medida que avanza la gestación, y esta disminución es aún más pronunciada en las mujeres con DG. Hacia el final del embarazo, la cantidad de IRS1, una de las moléculas de señalización del músculo esquelético, es menor en las mujeres embarazadas que en las no embarazadas. De hecho, en **las mujeres con DG** la disminución de los niveles de IRS1 y la

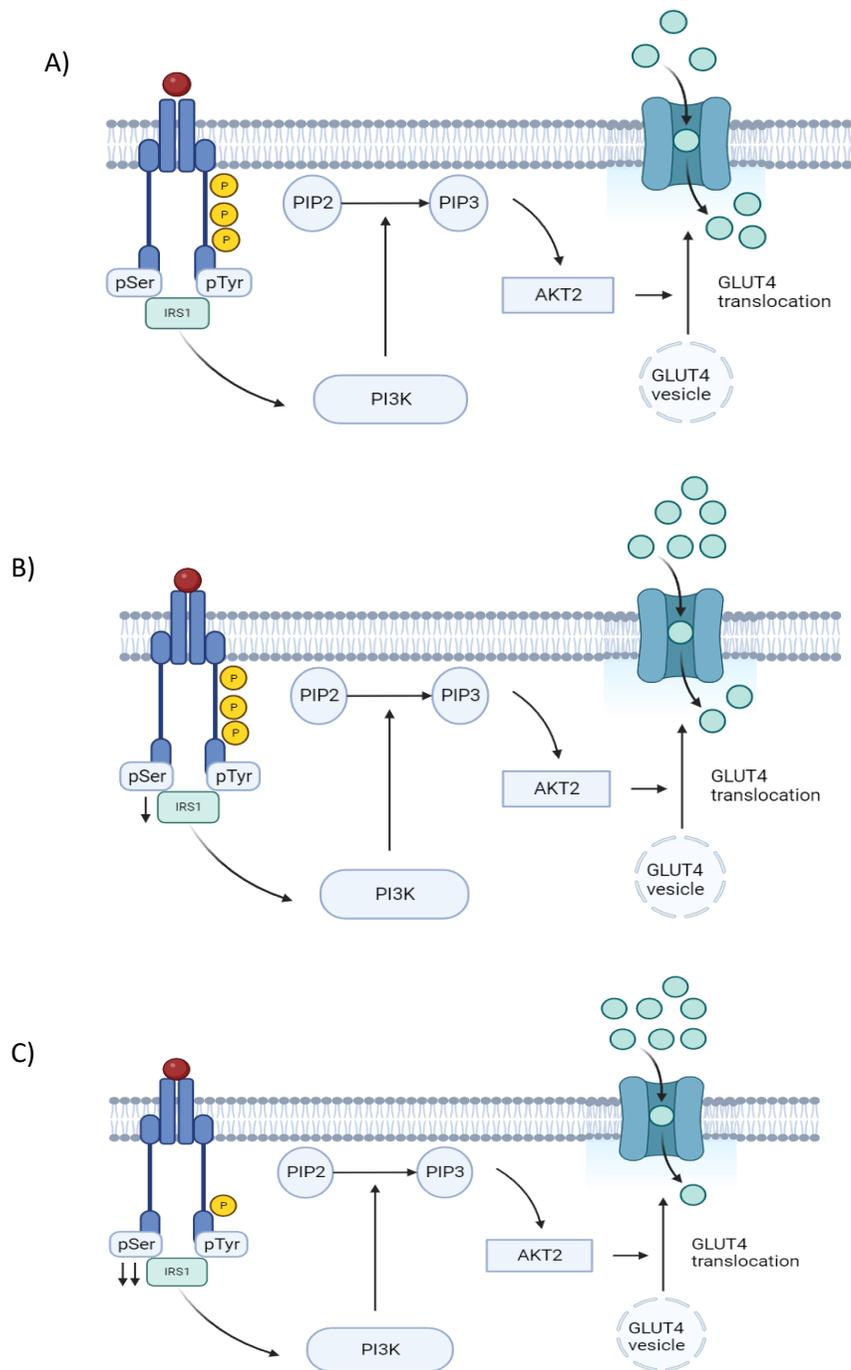


Figure 2- Mecanismo molecular de la glucosa. A) Situación normal, sin resistencia a la insulina. B) Situación durante el embarazo sin DG. C) Situación con DG. Bolas rojas: insulina, Bolas azules: glucosa.

autofosforilación de IR β es menor que en las embarazadas sin diabetes , lo

que se traduce en una captación de glucosa un 25% menor en el músculo esquelético (25,26). Además, en la DG se han observado modificaciones en la expresión y fosforilación de los reguladores de la señalización de la insulina, como IRS1, fosfatidilinositol 3-cinasa (PI3K) y GLUT4 (27) (Figura 2). Algunas de estas alteraciones moleculares persisten después del embarazo (28).

Consecuencias

Las complicaciones a corto plazo se pueden dividir en las siguientes (29):

- Nivel materno: complicaciones hipertensivas como la preeclampsia.
- Nivel obstétrico: mayores tasas de cesáreas, partos instrumentales, laceraciones del canal del parto.
- Nivel perinatal: macrosomía, polihidramnios, ictericia, hipoglucemia neonatal, y menos frecuente la muerte perinatal.

Durante mucho tiempo se debatió si los malos resultados asociados a la DG se debían a la hiperglucemia materna o a otros factores de riesgo que acompañan a menudo a estas mujeres (sobrepeso, aumento de la edad, la reducción de la actividad física o la pertenencia a una minoría étnica). Sin embargo, el estudio HAPO, relacionó claramente que la hiperglucemia materna aumenta de forma independiente y lineal el riesgo de complicaciones durante el embarazo y en la descendencia. El estudio reveló que los valores de glucosa plasmática en ayunas de la SOG se asociaron más fuertemente con malos resultados que los valores a la hora y dos horas de la prueba (30). Otros estudios, han revelado que el

tratamiento de la DG y el control de la glucosa durante el embarazo es eficaz para prevenir complicaciones maternas y fetales a corto plazo (31).

La principal consecuencia a largo plazo para la madre es el mayor riesgo de desarrollar DM2. Así, el riesgo de padecer DM2 en mujeres que hayan padecido DG puede ser hasta 7 veces mayor que en mujeres con embarazos normales (32). Además, otros factores como son un IMC elevado al inicio del embarazo, mayores niveles de glucosa en la sobrecarga de glucosa y la necesidad de insulina durante el embarazo aumenta el riesgo del desarrollo de DM2 en el futuro (33). El estudio HAPO, sobre mujeres con DG no tratadas desveló que el no tratamiento de estas pacientes conlleva importantes riesgos a largo plazo tanto para la madre como para el niño. En algunas poblaciones, el riesgo de diabetes tipo 1 (DM1) también aumenta tras una DG. Además, las mujeres con DG tienen mayor riesgo de padecer síndrome metabólico y enfermedades cardiovasculares, renales, hepáticas y retinianas (34) incluso de forma independiente a que desarrollen o no diabetes (35) .

Para la descendencia, la principal consecuencia es desarrollar obesidad y complicaciones metabólicas. Presentando un riesgo hasta 8 veces mayor de desarrollar prediabetes o diabetes en comparación con los controles (36). Igualmente, el riesgo de sobrepeso y síndrome metabólico está aumentado entre dos y cuatro veces respectivamente respecto la población de referencia, además la sensibilidad y la secreción a la insulina se ven reducidas (37,38). El estudio HAPO-FUS (seguimiento a largo plazo del estudio HAPO), confirmó que los niños de madres con DG tenían riesgo de presentar mayores hiperglucemias, resistencia a la insulina, adiposidad y riesgo cardiovascular. También sugirió que aunque la adiposidad materna

es un factor de riesgo para la obesidad en la descendencia, la DG se mantuvo como un factor de riesgo incluso ajustando por IMC materno (39).

Factores de riesgo

Entre los factores de riesgo de la DG se incluyen la edad materna avanzada, la etnia, la obesidad, los antecedentes de diabetes gestacional o macrosomía en gestaciones previas y los antecedentes familiares de DM2 (40) (Figura 3).

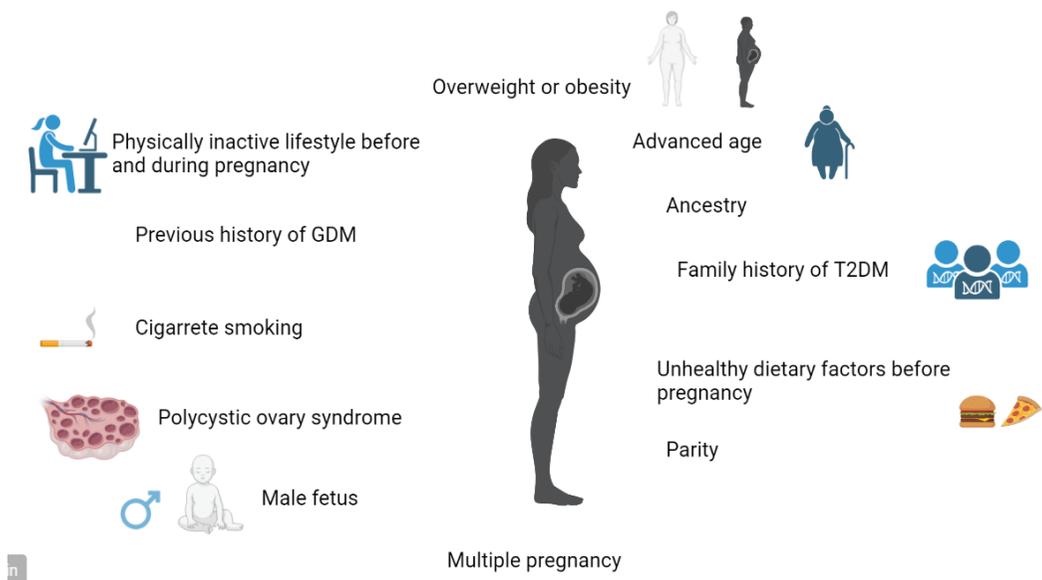


Figure 3- Factores de riesgo para el desarrollo de DG.

El sobrepeso u obesidad es el factor de riesgo más importante para desarrollar DG en mujeres embarazadas (41). Fumar durante el embarazo, así como la exposición pasiva al tabaquismo también aumenta el riesgo de DG, independientemente del IMC previo al embarazo y de otros factores de riesgo (42). Se ha demostrado que la práctica de actividad física antes y durante el embarazo reduce el riesgo de DG (43,44). Varios factores dietéticos podrían afectar al riesgo de DG, aunque no se pueden extraer

conclusiones definitivas (41). Hay pruebas de que los niveles bajos de vitamina C y vitamina D al principio del embarazo, así como el aumento de la ingesta de grasas en la dieta durante el embarazo, pueden aumentar el riesgo de DG. En general, se ha identificado que los alimentos conocidos como ultraprocesados pueden aumentar el riesgo de DG, independientemente de la grasa corporal y la actividad física (45–47). Por el contrario, consumir más fibra y frutos secos y seguir una dieta caracterizada por un alto consumo de fruta, verduras de hoja verde, aves y pescado parece ser beneficioso para reducir el riesgo de DG (48).

Son muchos los trabajos que han demostrado una asociación entre la edad y el riesgo de DG. Las mujeres mayores de 40 años demostraron tener más del doble de riesgo de desarrollar DG en comparación con las mujeres menores de 30 años (con una prevalencia del 9,8% frente al 4,1%, respectivamente). Este elevado riesgo persistió incluso después de tener en cuenta otros factores significativos que también influyen en la DG (44). Además, parece existir un mayor riesgo de DG entre las mujeres portadoras de fetos masculinos (49), y algunos estudios sugieren una mayor probabilidad de DG en los embarazos gemelares, aunque este hallazgo no es universalmente consistente (50).

En países con poblaciones diversas, como Australia, EE.UU. y Canadá, se han observado notables variaciones en la prevalencia de DG entre las distintas etnias (51). Las razones subyacentes de estas diferencias étnicas son probablemente complejas y abarcan factores como las variaciones en la adiposidad corporal, las elecciones de estilo de vida (incluidas la dieta y la actividad física) y la predisposición genética, entre otros.

Factores genéticos

Se sabe que la genética está implicada en la etiología de la DG, habiéndose asociado algunos genes con el desarrollo de la misma, pero los estudios al respecto son escasos y limitados (52). De hecho, hay pocos estudios de asociación de genoma completo (GWAS) relacionados con la DG, y los que hay, se han realizado en pocas etnias. Alrededor del 5% de las mujeres blancas con DG presentan variantes monogénicas en el gen de la glucoquinasa (*GCK*) (53). Un estudio realizado en población china identificó 23 polimorfismos de nucleótido único (SNP, de sus siglas en inglés) relacionados con la DG, los cuales se encontraban en 4 genes: *SYNPR*, *CDH18*, *CTIF* y *PTGIS* (54). Hay dos estudios realizados en mujeres coreanas, uno de ellos identificó algunos polimorfismos en los genes *CDKAL1* y *MTNR1B* (55), y el otro encontró cómo algunas variantes poligénicas asociadas con la DG influían en el riesgo de DM2 más adelante en la vida (56). Recientemente, algunos estudios han relacionado las variantes genéticas de la DG con la DM2 (57). Hay otros genes que están estrechamente relacionados con el metabolismo de los hidratos de carbono y la insulina, como *INS*, *IRS1*, *GCK*, *KCNJ11*, *TCF7L2*, *MTNR1B*, *CDKAL1*, *IGF2BP2* y *KCNQ1* que podrían estar implicados en la fisiopatología de la DG (58).

Epigenética

La epigenética abarca los cambios estructurales y bioquímicos hereditarios que no alteran la secuencia de ADN. Las alteraciones epigenéticas pueden estar influidas por el ambiente y ser heredadas por la descendencia. Los mecanismos epigenéticos están involucrados en diversos procesos fisiológicos y patológicos a través de la regulación de la expresión de ciertos

genes, pudiendo alterar desde la accesibilidad de las polimerasas a la localización de la cromatina (59).

Mecanismos epigenéticos

Existen distintos tipos de modificaciones epigenéticas, entre las más estudiadas se encuentran: la metilación del ADN, la modificación de histonas y los non-coding RNA (nc-RNA), o en español ARN no codificante (Figura 4).

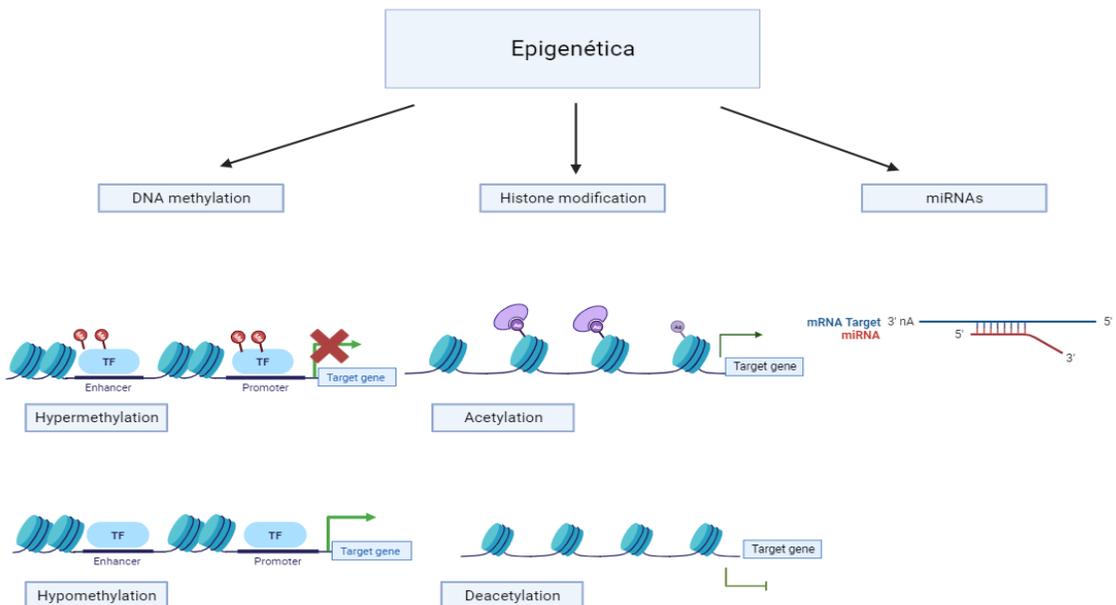


Figure 4- Tipos de modificaciones epigenéticas.

Metilación del ADN

La metilación del ADN es el mecanismo epigenético más estudiado. El patrón de metilación más común resulta de la transferencia de un grupo metilo a la posición C5 de la citosina dentro de los dinucleótidos CpGs. Este proceso se lleva a cabo por un grupo de ADN metiltransferasas. Se conocen como islas CpGs a tramos cortos de ADN con un contenido alto de GC y una mayor frecuencia de dinucleótidos CpG en comparación con el resto del genoma. La localización de los CpGs se pueden referenciar teniendo en

cuenta su posición en el genoma, o teniendo en cuenta su posición con respecto a la isla (60).

Según el genoma encontramos las siguientes posiciones:

- Cuerpo del gen
- Zonas intergénicas
- Región 3' no traducida (UTR)
- Zonas promotoras: 5'UTR, o sitios de inicios de la transcripción (TSS, en inglés) que pueden estar localizados a 200 o 1500 nucleótidos aguas arribas del gen. El 60% de los CpGs se encuentran en estas regiones.

Según la localización de la metilación del sitio CpG con respecto a las islas CpGs existen las siguientes localizaciones (61):

- *Shore*: adyacente a islas CpGs, hasta 2kb de la isla CpG.
- *Island*: dentro de la isla de CpGs
- *Shelf*: De 2 a 4 kb de la isla CpG.
- *Open-sea*: regiones alejadas de las islas CpGs, a más de 4 kb.

La metilación de ADN juega un papel muy importante en la regulación de procesos tanto fisiológicos como patológicos, y la metilación errónea del ADN se asocia a menudo con el desarrollo de múltiples enfermedades (62). El proceso por el cual la metilación del ADN controla la expresión génica es mediante la regulación dinámica del estado de las islas CpGs en las regiones reguladoras del gen. Por lo general, una hipermetilación en una isla CpG en la región reguladora del gen suele estar asociado con la compactación de la cromatina resultando en el silenciamiento del gen. Sin embargo, una hipometilación deja una apertura más laxa de la cromatina

permitiendo la transcripción del gen (Figura 4). Se conoce que la metilación del ADN está implicada en la modulación de varios procesos biológicos como en el desarrollo embrionario, la impronta genómica, la inactivación del cromosoma X, la diferenciación celular y la proliferación. Por otro lado, la metilación errónea en algunas regiones promotoras en genes estrechamente relacionados con tumores, puede resultar en silenciamiento de genes supresores de tumores o en la activación de genes oncogénicos, dando lugar a células cancerosas (59,63).

Modificación de histonas

Otro mecanismo epigenético importante es la modificación de la cromatina ya que juega un papel importante en el control de la transcripción génica. La remodelación de la cromatina juega un papel importante en la regulación de cómo los genes se expresan, y de lo que ocurre tras las modificaciones químicas en las proteínas de las histonas que se producen después de su síntesis. Los cambios químicos que experimentan, como la acetilación y la metilación, se producen generalmente en los residuos específicos de lisina en las colas de histona centrales (64). Estas modificaciones en las histonas otorgan una identidad única a los nucleosomas, influyendo en cómo se activan o se desactivan los genes durante diferentes etapas biológicas. Una acetilación de histonas suele estar relacionada con una activación de la transcripción génica, mientras una desacetilación se asocia a un silenciamiento génico (Figura 4). Los mecanismos por los cuales las modificaciones de las histonas regulan la transcripción génica han sido bien documentados. Los cambios en la estructura de la cromatina causados por las modificaciones de las histonas pueden cambiar la conformación espacial de la polimerasa de ADN, lo que lleva a cambios en la accesibilidad de los factores de

transcripción clave, alterando posteriormente la actividad transcripcional. Existen distintas enzimas que están implicadas en el proceso de la modificación de histonas como son:

- Las histonas acetiltransferasa (HAT).
- Las desacetilasas de histonas (HDAC).
- Las histonas metiltransferasas (HMT).

Estas enzimas intervienen tanto en la regulación de la modificación de histonas como en el mantenimiento de la configuración de la cromatina (65,66).

ARN no codificantes

Otra forma de modificación epigenética es la mediada por los nc-RNA. Estos ARN presentan un papel fundamental en el control de la expresión génica y la modificación de la cromatina. Los nc-RNA pueden interactuar con el ADN, ARN mensajero y proteínas para influir en la actividad de manera reversible y heredable. Se ha visto que este mecanismo induce silenciamiento génico mediado por ARN a través del mecanismo de metilación de ADN dirigida por ARN, lo que indica que los nc-RNA pueden participar en la regulación de otros patrones de modificación epigenéticas (67).

Existen distintos tipos de nc-RNA y se pueden dividir en dos categorías según el tamaño: pequeñas cadenas no codificantes y largas cadenas no codificantes (ln-RNAs). Entre los de pequeña cadena se encuentran los siRNA, los miRNA y los piRNA (68,69). A continuación, se describen los distintos tipos de nc-RNA:

- miRNA: son RNA cortos de cadena simple, generalmente de unos 20-22 nucleótidos de longitud, que se encuentran principalmente en regiones cromosómicas propensas a cambios estructurales. Se unen a secuencias específicas de ARNm y silencian la expresión del gen. Esto lo logra a través de la inhibición de la traducción. Los miRNA están involucrados en gran cantidad de procesos biológicos y pueden actuar como represores de genes (70,71).
- SiRNA: son RNA cortos, proveniente de RNA de doble cadena que derivan de la replicación de virus, actividad de los trasposones o de la transcripción génica. Suele intervenir en los procesos de regulación transcripcional de forma parecida a los miRNA. Estos pequeños fragmentos de RNA ejercen su función al unirse con las proteínas argonautas (AGO) (72,73).
- piRNA: molécula de RNA de unos 26-31 nucleótidos de longitud. El nombre piRNA (Piwi-interacting RNA), refleja que estas moléculas se unen a las proteínas Piwi en condiciones fisiológicas. Las proteínas Piwi silencian los genes implicados en la regulación de otros genes implicados en el desarrollo (homeobox) uniéndose a los elementos de respuesta del complejo de reguladores del grupo Polycomb. Se ha visto que la interacción de los piRNA tiene un papel importante en la regulación epigenética (74).
- InRNA: los InARN son de más de 200 nucleótidos y no codifican proteínas. Tienen diversas funciones de regulación actuando como moduladores de la cromatina, influyendo en la estructura de su estructura y la accesibilidad del ADN a las maquinarias de transcripción y regulación génica (75).

Tecnología para la detección de metilación del ADN

La metilación de ADN es la marca epigenética más estudiada gracias a los avances tecnológicos para medir esta modificación. Esto ha sido posible debido a la estabilidad química y temporal de la metilación de ADN. Químicamente, la metilación de ADN es una modificación covalente con una estabilidad química relativamente alta debido a la unión covalente del metilo a los dinucleótidos de CpGs. En los últimos años las técnicas para la detección de la metilación de ADN han avanzado mucho. Cada una de las técnicas responden a distintas necesidades de la investigación que se lleve a cabo (Figura 5). Además, a diferencia de la modificación de las histonas, la metilación y estructura no se pierde durante el proceso de extracción de ADN (76,77).

Técnicas basadas en la conversión con bisulfito

Un paso fundamental para la detección de la metilación del ADN es la conversión a bisulfito. Este método está basado en el tratamiento del ADN genómico con bisulfito de sodio, lo que resulta en la desaminación de las citosinas no metiladas a uracilo, dejando intactas las citosinas metiladas. Esto permite distinguir entre citosinas metiladas (5mCs) y citosinas no metiladas en todo el ADN genómico. El *gold standar* del estudio de la metilación de ADN es la secuenciación de ADN bisulfitado, que proporciona un enfoque altamente cuantitativo y eficiente para identificar 5mC con resolución a nivel de un solo par de bases (78).

Dependiendo del enfoque de cada estudio, será conveniente emplear una técnica u otra. En el caso de querer saber cómo se encuentra la metilación de un locus específico existen distintas técnicas como la secuenciación de

PCR específica de metilación y la pirosecuenciación. En la PCR específica de metilación, el ADN convertido en bisulfito se utiliza como molde para procesos posteriores de PCR determinando así el estado de metilación en los loci específico de interés mediante primers diseñados específicamente

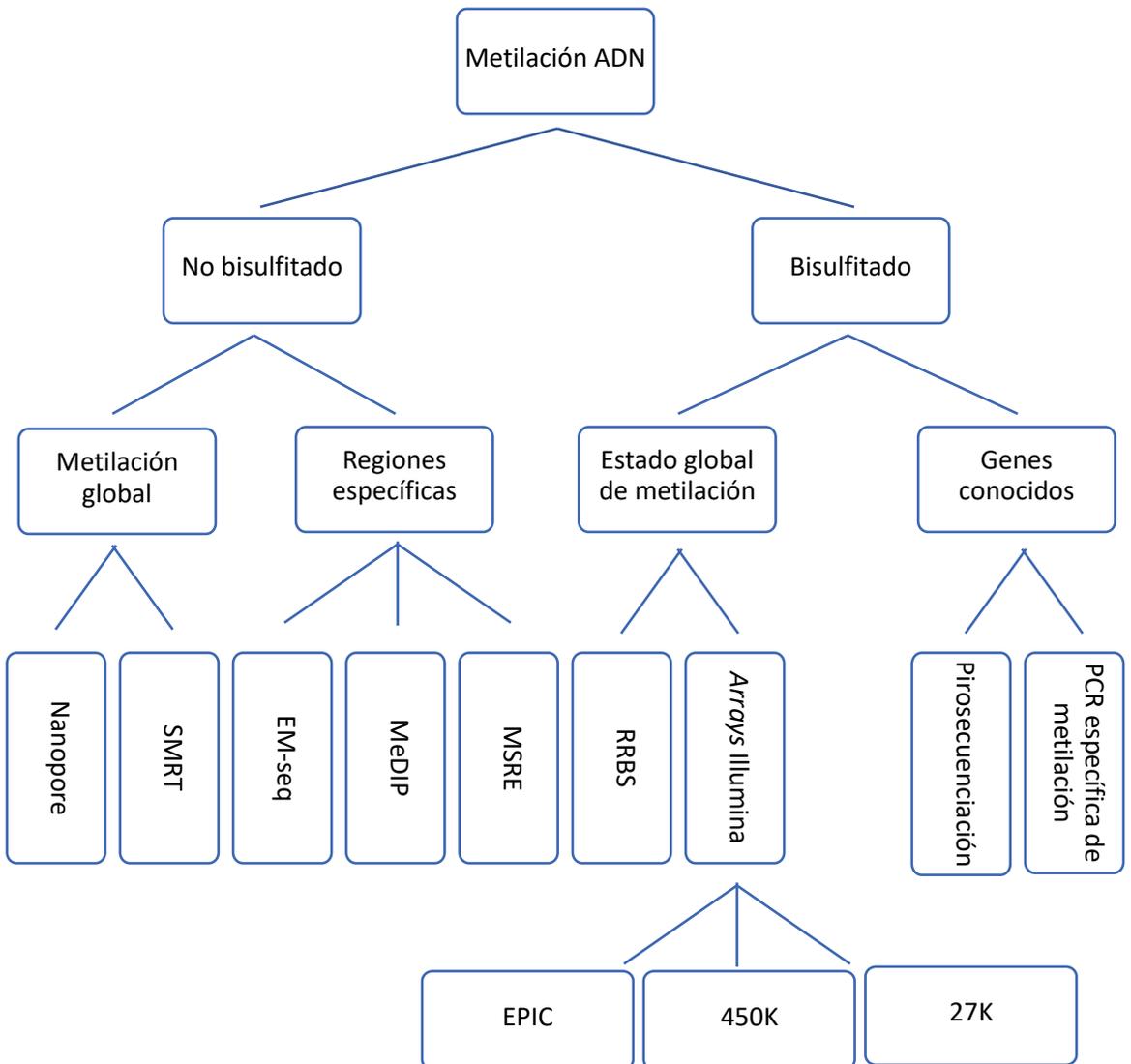


Figure 5- Tipos de tecnología para el estudio de la metilación de ADN. SMRT: tecnologías de secuenciación a tiempo real, EM-seq: metilación por secuenciación enzimática, MeDIP: inmunoprecipitación del ADN metilado, MSRE: digestión con enzimas de restricción sensibles a la metilación, RRBS: secuenciación por bisulfito de representación reducida.

para reconocer el molde de ADN metilado o no metilado (79). Por otro lado, la pirosecuenciación es un método basado en la secuenciación que puede monitorizar cuantitativamente la incorporación de nucleótidos en tiempo real mediante la detección de luminiscencia de pirofosfato. La pirosecuenciación puede proporcionar una cuantificación reproducible y precisa del estado de metilación del gen de interés con una alta resolución cuantitativa (80).

Por otro lado, el interés creciente de conocer el metiloma completo del ADN ha ido incrementando. En los últimos 10 años se ha ido desarrollando la tecnología necesaria para conocer el perfil epigenético de grandes regiones del genoma. Estos estudios se conocen como *Epigenome-wide association study* (EWAS). Entre los métodos utilizados para la detección del ADN metilado, los *microarrays* o también llamados *arrays* de Illumina son los más usados.

A lo largo de los años han ido apareciendo distintos tipos de arrays. En los inicios, se usaba el Illumina 27k, que contenía 27.570 sitios CpG distribuidos entre 14.495 genes (81). Posteriormente apareció el Illumina 450k que contenía más de 485.000 regiones de metilación, cubriendo también regiones promotoras (82). Hasta hace poco era el tipo de array más utilizado en la comunidad científica. Recientemente Illumina sacó EPIC, otro array que en esta ocasión cubría hasta un total de 868.564 CpGs a lo largo del genoma. Este nuevo chip incluye >90% de las sondas del 450K y añade nuevas islas CpG, sitios CpGs fuera de islas, así como regiones enhancers y sitios de unión a factores de transcripción basados en los proyectos FANTOM y ENCODE (referencia de evaluación del EPIC). Además

cubre regiones promotoras, y regiones codificadoras de genes relacionadas con el cáncer, el desarrollo y enfermedades complejas (83). En los últimos años ha habido un crecimiento de estudios que usan este tipo de tecnología. Aunque han supuesto un gran avance, y es la técnica más utilizada, representa un porcentaje pequeño de todos los sitios CpGs del genoma humano.

Cuando solo se quiere ver una pequeña fracción de la metilación del genoma, se han desarrollado estrategias de enriquecimiento, basadas en enzimas de restricción, como la secuenciación por bisulfito de representación reducida (RRBS). RRBS ha sido aplicado en análisis del metiloma de varias especies incluidos no mamíferos como pescados, insectos y plantas (84).

Otras técnicas no dependientes de bisulfitación

Aunque la bisulfitación del ADN se considere el *gold standard* para la detección de la metilación, no siempre se necesita de este paso. Existen alternativas basadas en la restricción enzimática o enriquecimiento por afinidad como la digestión con enzimas de restricción sensibles a la metilación (MSRE) y la inmunoprecipitación del ADN metilado (MeDIP), seguida de una amplificación por sonda dependiente de ligadura múltiple específica de metilo (MS-MLPA). La MSRE depende de la secuencia, lo que limita el análisis de la metilación a regiones que contienen los sitios de restricción. Otro método que no usa bisulfitación es la metilación por secuenciación enzimática (EM-seq), que solo aplica pasos enzimáticos para desaminar la citosina, mientras que la 5mC permanece inalterada.

Recientemente, debido al desarrollo de la secuenciación con lecturas de larga cadena, se han empezado a utilizar tecnologías de secuenciación a

tiempo real (SMRT, del inglés) de PacificBiosciences (PacBio) (85) y la secuenciación por nano poros de Oxford Nanopore Technologies (ONT) (86). Ambas tecnologías pueden secuenciar ADN nativo, por tanto, se evita la degradación del ADN debida a la conversión de bisulfito.

Procesamiento y finalidad de los estudios EWAS

El principal objetivo de un estudio de EWAS es encontrar variantes epigenéticas que se asocien a algún fenotipo de interés (87). Las modificaciones de la metilación del ADN pueden ser un mecanismo que expliquen la relación entre la exposición temprana y posteriores fenotipos metabólicos en la vida. Pero además de tener un papel mecanístico, la metilación del ADN puede servir como biomarcador de la exposición o el fenotipo (88).

Flujo de trabajo en el análisis bioinformático de los EWAS

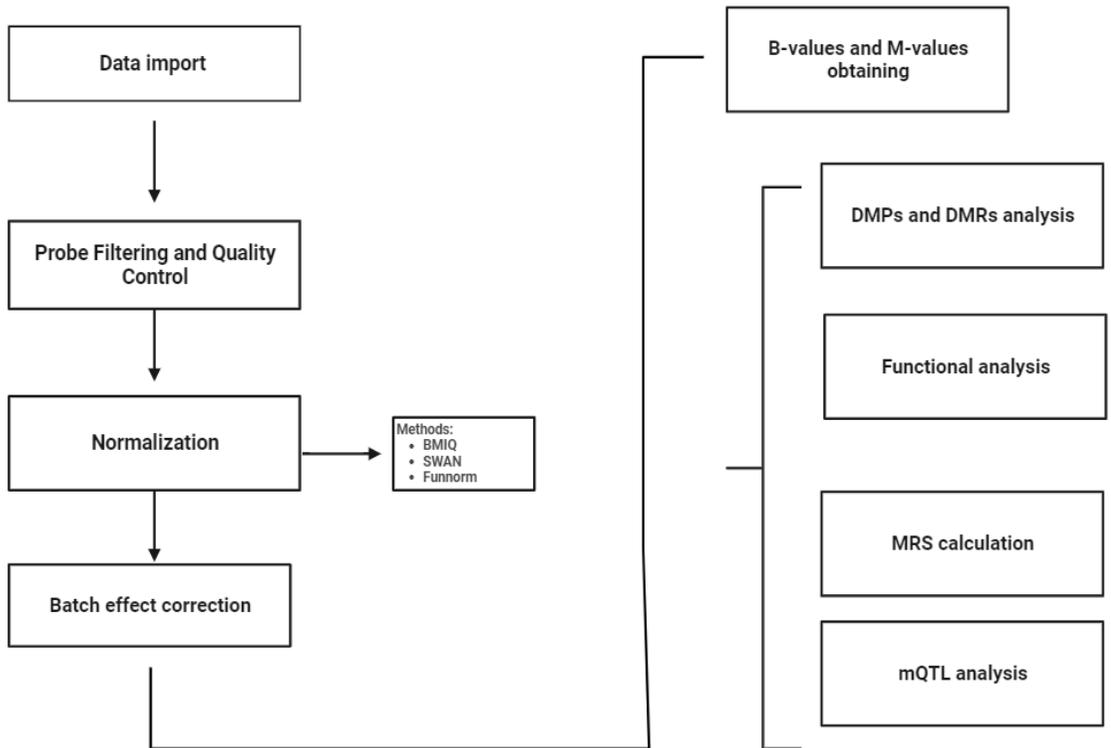


Figure 6- Flujo de trabajo para el análisis y obtención de resultados de los arrays de metilación de Illumina.

El flujo de trabajo establecido y los paquetes de R (bioconductor) que se utilizan de forma general aparece en la figura 6. Los pasos a seguir incluyen obtención y carga de los datos, control de calidad, normalización, exploración de los datos, filtrado de las sondas, análisis de las posiciones/sitios diferencialmente metilados (DMPs, del inglés) y análisis de las regiones diferencialmente metiladas (DMRs, del inglés) (89).

Obtención y carga de los datos

Los datos de metilación que se obtienen de Illumina son los archivos Intensity Data (IDAT). Almacena un resumen de las intensidades de cada

sonda del array. Para importar los datos de metilación en R se utiliza el paquete `minfi`, el cual necesita los IDAT y una hoja con las muestras. Una vez tenemos la información cargada en R sobre las muestras y su localización, se puede leer la intensidad de los datos a partir de los archivos IDAT. Se crea un objeto que contiene todos los datos de intensidad brutos, de los canales de color rojo y verde, para cada una de las muestras.

Control de calidad

Una vez obtenido los datos, se procede a evaluar su calidad. Para ver la calidad de la señal de cada CpG se calcula los valores P de la detección de las sondas. `minfi` lo calcula comparando la señal total (Metilado + No-metilado) para cada sonda comparada con el nivel de señal de fondo. A menor valor de P mejor es la calidad de la señal, un $p > 0.01$ se suele considerar una baja calidad. El p-valor medio para cada muestra permite tener una idea de la calidad global del array.

Normalización

Los arrays de Illumina usan dos tipos de hibridación química distinta para medir la metilación en las sondas, tipo I y tipo II, que producen diferentes distribuciones de valores beta. Estas distribuciones deben normalizarse para evitar la detección sesgada de DMPs enriquecidas para sondas de tipo I. Existen distintos tipos de normalización: normalización por cuantiles de mezcla beta (BMIQ de sus siglas en inglés), normalización por cuantiles de subconjuntos dentro de microarrays (SWAN del inglés), la corrección basada en picos (PBC) o la normalización funcional (Funnorm) (90–93). Este último método de normalización, Funnorm, es el que produce los resultados más replicables en los análisis donde se esperan cambios de la metilación global, además de eliminar gran cantidad de variación técnica

en el proceso de normalización no supervisada, lo que mejora la corrección posterior del efecto técnico. Sin embargo, la técnica más ampliamente utilizada es BMIQ, usada por defecto cuando se utiliza el paquete ChAMP.

Filtrado de las sondas

Las sondas con bajo rendimiento y calidad suelen filtrarse antes de seguir con los análisis. En general se filtran las sondas que han fallado en una o más muestras basándonos en el valor p de detección obtenido en el control de calidad. Dependiendo de la naturaleza de los datos, y el diseño del estudio se puede elegir la eliminación de sondas localizadas en los cromosomas sexuales (X e Y), también se suelen eliminar aquellas sondas que se localizan en regiones de polimorfismos de nucleótido único (SNPs del inglés). También se pueden filtrar sondas que hayan demostrado reactividad cruzada, en otras palabras, que se haya visto que mapean en múltiples lugares del genoma.

Eliminación de la variación debida a la técnica

En un experimento ideal, todas las muestras se procesarían en un solo lote. Sin embargo, esto es imposible para tamaño de muestras muy grandes. La naturaleza de estudios longitudinales implica que tienen un mayor riesgo de variación debido a la técnica o *Batcheffects*, ya que las muestras se recopilan y procesan en diferente momento.

El método utilizado por ChAMP para la identificación del *batcheffect* es la descomposición en los valores singulares (SVD, por sus siglas en inglés). SVD correlaciona los componentes principales con la variación biológica y técnica, utilizando información de la hoja de muestras proporcionada (94). Para la corrección de este efecto se suele utilizar ComBat, que utiliza un método de Bayes empírico para corregir los *batcheffects* (95). Este método,

necesita dos entradas: por un lado, la variación biológica causada por la variable de interés y por otro lado las variables del batcheffect que queremos eliminar (el array, la placa...), identificado previamente mediante la técnica SVD.

Tipos de análisis de EWAS

Tras el procesamiento de los datos se procede a extraer los valores M (M-values) y los valores β (β -values). El nivel de metilación de un sitio CpG se puede expresar en forma de valores β o M. Los β se calculan mediante la fórmula que se expresa a continuación y oscilan entre 0 y 1. Los valores M se calculan mediante la transformación logarítmica de los β -value (96). Los β -value se usan generalmente para las representaciones gráficas, mientras los M-values suelen usarse para los análisis estadísticos, aunque no existe un consenso sobre esto y ambos son válidos.

Obtención de las DMPs

El objetivo principal de este tipo de estudios es encontrar diferencias en los niveles de metilación entre los distintos grupos. Para ello, a partir de la matriz de M-values o β -values, se usa el paquete limma de bioconductor para obtener los DMPs.

El paquete limma realiza una comparación por pares de valores beta entre grupos, ajustando el mismo modelo lineal general a cada sonda por separado y calculando un estadístico t moderado y un valor p no ajustado. Un estadístico t moderado es la relación entre el valor beta de un CpG y un error estándar (SE) agrupado. Al agrupar la información de todos los CpG, limma modera el SE en sondas individuales para mejorar la inferencia sobre cada CpG.

Los valores p obtenidos deben ser ajustados para comparaciones múltiples utilizando distintos métodos entre los que se encuentra Bonferroni y Bejnamini y Hochberg que nos devuelve un valor de False Discovery Ratio (FDR). Normalmente se utiliza un FDR del 5%, asumiéndose que el 5% de los DMPs significativos pueden ser falsos descubrimiento. Este umbral puede ser modificado asumiendo las consecuencias del mismo. El estadístico utilizado por limma, permite incluir covariables y hacer ajustes del modelo.

Aunque el t-test de limma sea el comúnmente más utilizado, dependiendo del tipo de estudio o de los ajustes según la normalidad de los datos se pueden hacer distintas pruebas estadísticas para detectar los DMPs.

Obtención de las DMRs

Las DMRs son regiones genómicas compuestas por varios DMPs contiguos. A menudo están asociadas con una región génica específica, como las islas CpG en regiones promotoras, pudiendo encontrarse también en regiones intergénicas. Existen múltiples definiciones, enfoques y herramientas bioinformáticas para su identificación. La definición más común para una DMR es que en una región de entre 500-2000pb contenga al menos 2 DMPs. Dependiendo de la región a estudiar la ventana considerada como DMR se puede modificar para incluir toda la zona reguladora como sitios TSS, 5'UTR y el primer exón (97).

Existen distintas herramientas para analizar las DMRs. La mayoría dependen de definir el umbral de DMPs, la distancia genómica entre DMPs y el número mínimo de DMPs por DMR. La herramienta más ampliamente usada es Bumhunter que detecta los DMRs a través de la "búsqueda de

picos” o “detección de crestas”. Para cada sitio CpG, Bumhunter produce una pendiente a partir de un modelo lineal basado en el fenotipo. La pendiente suavizada se traza y se analiza en búsqueda de “crestas” que pasen un umbral establecido (92). Otra herramienta más actual es DMRcate. Realiza una regresión a nivel de metilación en cada sitio CpG basado en el fenotipo, teniendo en cuenta las covariables. Esto es seguido por un suavizado gaussiano (promedio de efecto) y la agrupación de sitios CpGs cercanos según una ventana definida por el usuario (82). DMRcate es mucho más demandante en la computación pero supera a nivel de potencia y precisión a Bumhunter (98), por lo que a día de hoy es la que más se utiliza. También existe otra herramienta utilizada popularmente para la identificación de DMRs llamada Comb-p (99). No es específica para metilación, pero su rendimiento es comparable al DMRcate. Comb-p analiza y corrige los valores p en una ventana o región genómica definida por el usuario basada en sondas vecinas ponderadas, y luego se asigna un valor p a toda la región genómica.

Estudios de enriquecimiento funcional

Una vez que se obtienen los DMPs y/o DMR lo interesante es saber en qué mecanismos influyen los genes en los que se encuentran. Existen muchas formas de realizar este tipo de estudios.

La ontología de genes (GO, del inglés) es de gran interés para conocer las posibles funciones que tiene los genes encontrados. Las ontologías suelen consistir en un conjunto de términos o conceptos con relaciones que operan entre ellas. La GO describe el conocimiento del ámbito biológico en relación con tres aspectos: función molecular, componente celular y proceso biológico. En resumen, las GO nos da información sobre la función

del producto génico (100). Para obtener los GO a partir de los DMPs obtenidos, el paquete `missMethyl` de R tiene el algoritmo `GOMeth` que devuelve los términos GO con su nivel de significancia y los genes que están relacionados con ese término. Otro dato de interés es conocer las rutas de interés implicadas en las regiones y posiciones diferencialmente metilada. Para ello, `missMethyl` también tiene una función para extraer las rutas biológicas asociadas a los sitios de interés a partir de Kyoto Encyclopedia of Genes and Genomes (KEGG). KEGG es una base de datos para comprender las funciones y utilidades de alto nivel del sistema biológico, especialmente de datos moleculares a gran escala generados por secuenciación del genoma y otras tecnologías experimentales de alto rendimiento (101). También a partir de la lista de genes anotados en los CpGs que se obtienen se pueden utilizar otras herramientas como Enrichr (102), ReviGo (103), Gene Ontology website (104) entre otros.

Otra aproximación que se suele llevar a cabo es conocer como interaccionan las proteínas resultantes de los genes en los que hemos encontrados los DMPs y/o DMRs. Esto puede dar información de qué mecanismos pueden estar influyendo en el fenotipo de interés y arrojar luz a la interpretación biológica de los resultados. Para este tipo de estudios se usan distintas herramientas disponibles tanto en web o software específico como en determinados paquetes de R. Entre estas herramientas encontramos por un lado STRING data-base, Cytoscape y David data-base (105–107):

- STRING es una base de datos que predice la interacción proteína proteína (PPI, del inglés). Las interacciones pueden ser tanto físicas como funcionales.

- Cytoscape permite la visualización de redes de interacciones compleja y mediante aplicaciones específicas permite agrupar las interacciones de proteínas en función de su funcionalidad.
- DAVID debe sus siglas en inglés Base de Datos de Anotación y Descubrimiento Integrado que proporciona un amplio conjunto de herramientas de anotación funcional para que los investigadores comprendan el significado biológico de grandes listas de genes.

El uso correcto de estas herramientas nos puede ayudar a identificar posibles mecanismos implicados en el fenotipo o patología de interés.

Cálculo de puntuación del riesgo de metilación o methylation risk score

El creciente interés en encontrar marcadores de riesgo de enfermedades poligénicas ha llevado a la creación de los scores de riesgo poligénicos (GRS), los cuales permiten cuantificar el riesgo acumulado de una enfermedad causado por variantes con tamaños de efectos pequeños en múltiples genes identificados en los estudios de Genome-wide association (GWAS). Del mismo modo, la puntuación de riesgo de metilación o methylation risk score (MRS) combina los DMPs identificados en EWAS para encontrar indicadores significativos de riesgo a nivel individual.

Para la construcción de un MRS en primer lugar hay que identificar los DMPs asociados a un fenotipo de interés. El efecto de cada DMP puede ponderarse en función del tamaño de su efecto. La figura 7 refleja cómo se calcula el MRS. Se han aplicado con éxito MRS en diversas enfermedades complejas, tales como el cáncer y enfermedades metabólicas (108).

$$MRS_j = \sum w_i * \beta_i$$

w_i = el peso asignado a DMP.

β_i = nivel de metilación (en valores beta o M) en el DMP_i de cada individuo j.

Figure 7- Fórmula para el cálculo de MRS. Siendo w el peso del efecto de un CpG para un fenotipo de interés, y beta el valor beta o M del CpG de cada individuo.

Análisis de loci de rasgos cuantitativo de metilación (m-QTLs)

Los loci de rasgos cuantitativos (QTL, por sus siglas en inglés) son loci genómicos que influyen en un rasgo o fenotipo cuantificable. Hacen referencia a una región específica de los cromosomas que alberga uno o varios genes que controlan dichos rasgos (109). Los QTL se pueden utilizar para mapear la relación entre los niveles de metilación y el genotipo en un locus específico, y a esto se le conoce como loci de rasgos cuantitativos de metilación comúnmente abreviados como methQTL, metQTL o mQTLs. En otras palabras, los mQTL son genotipos individuales en loci específicos que pueden dar lugar a diferentes patrones de metilación de regiones genómicas extensas. Los mQTLs se han identificado de manera consistente en diversas enfermedades, poblaciones y etapas del desarrollo (110).

Existen varias formas de identificar mQTLs, un primer enfoque simple es a partir de los DMPs obtenidos asociados con nuestro fenotipo de interés. Se seleccionan SNPs con una frecuencia alélica menor (MAF, de sus siglas en inglés) ≥ 0.05 en la misma región genómica. Para ello existen softwares específicos como GoDMC, donde poder buscar los SNPs cercanos al CpG de interés. GoDMC es una herramienta que expone los resultados de un análisis a gran escala, donde se mapeó las influencias genéticas tanto en posición cis como trans en la metilación utilizando 36 cohortes con un total

de 27.150 muestras europeas de sangre (111). Los datos obtenidos se pueden consultar en la página web: <http://mqtl.db.godmc.org.uk/index>.

Existen paquetes disponibles para cálculos de mQTL más avanzados. Por ejemplo, el paquete R GEM (112) que puede modelar la relación compleja entre los niveles de metilación, el genotipo, las covariables y una variable ambiental, que podría reemplazarse por una variable fenotipo (112).

Toda esta aplicabilidad de los estudios EWAS, hace que sea una importante herramienta, útil a la hora de elucidar cómo algunas **enfermedades poligénicas** y/o multifactoriales como son la **diabetes** y la **obesidad** se desarrollan, buscando potenciales rutas metabólicas afectadas y **biomarcadores** de riesgo. Además, se conoce que el **epigenoma** es más cambiante durante los 1000 primeros días de vida, por lo que la DG nos ofrece un marco idóneo donde estudiar el efecto de un ambiente adverso sobre la salud de la madre y su descendencia.

Hipótesis y objetivos

Hipótesis

Nuestra hipótesis plantea que la exposición a la hiperglucemia, propia de la diabetes gestacional, causa modificaciones epigenéticas tanto en la madre como en el niño, que alteran la expresión de genes implicados en rutas metabólicas relacionadas con el desarrollo de enfermedades como la obesidad y la diabetes, y podrían explicar el mayor riesgo de estos sujetos (tanto de las madres como de la descendencia) a desarrollar patologías en etapas posteriores de la vida.

Objetivos

Para comprobar nuestra hipótesis, se propusieron los siguientes objetivos:

- Objetivo 1: Estudiar los patrones de metilación del ADN en muestras de sangre periférica de mujeres gestantes con y sin DG, tanto en el momento del diagnóstico como al final del embarazo. Dentro de este objetivo nos planteamos los siguientes objetivos específicos:
 - Objetivo 1.1: Identificar marcas epigenéticas involucradas en mecanismos o rutas relacionadas con la DG.
 - Objetivo 1.2: Identificar potenciales biomarcadores epigenéticos del desarrollo de DG, que puedan servir para una detección precoz.
- Objetivo 2: Comparar el perfil epigenético en muestras de saliva de la descendencia nacida de madres con DG con el de descendencia de madres sin DG.

- Objetivo 2.1: Relacionar los perfiles de metilación del ADN de la descendencia con el patrón de crecimiento del niño, para identificar posibles biomarcadores de riesgo de obesidad en el futuro.
- Objetivo 3: Evaluar la relación de un MRS de DM2 con la DG en distintas poblaciones y criterios diagnósticos.
 - Calcular un MRS basado en CpGs descubiertos en un EWAS de DM2 en cinco cohortes europeas
 - Evaluar si el MRS obtenido se asocia con la DG en nuestra población de estudio y en una cohorte independiente de validación con diferentes ancestros y criterios diagnósticos.

Materiales y métodos

Contexto

La presente tesis forma parte del proyecto de investigación “Efecto de la diabetes gestacional sobre el epigenoma de la madre y de la descendencia en los 1000 primeros días. Identificación de potenciales biomarcadores epigenéticos de riesgo de diabetes y obesidad” financiado por el ISCIII (PI18-01175), cuyo objetivo principal es crear una cohorte de mujeres embarazadas expuestas y no expuestas a la Diabetes Gestacional (cohorte EPIDG) y seguirlas a lo largo de los años, junto con su descendencia, con el fin de conocer mejor el efecto de la DG durante el embarazo sobre fenotipos metabólicos en el futuro, tanto en la madres como en los niños, e intentar buscar estrategias que frenen el ciclo obesidad/diabetes.

Los resultados que aquí se exponen corresponden a la fase de screening mediante arrays de metilación del ADN, que se ha realizado en un subgrupo de la cohorte EPIDG.

Sujetos del estudio

Las participantes del estudio forman parte de la cohorte EPIDG, un estudio de cohorte poblacional prospectivo que comenzó a principios de 2019 y aún sigue el reclutamiento y seguimiento de las pacientes. El reclutamiento tiene lugar en la Unidad de Diabetes y Embarazo del Hospital Universitario Virgen de la Victoria, Málaga, España. Los criterios de **inclusión** son los siguientes:

- Test O’Sullivan ≥ 7.8 mmol/l o 140mg/dl.
- Edad de 18 a 45 años.
- IMC < 40 .

Criterios de **exclusión**:

- Diagnóstico de DG en primer trimestre (EG<14 semanas).
- Alteraciones psicopatológicas que no garanticen adherencia al seguimiento.
- Pacientes que no puedan acudir con la frecuencia prevista a consulta programada.
- Gestaciones gemelares.
- Vómitos o intolerancia para la realización del test de tolerancia a la glucosa.

De todas las participantes se seleccionaron 32 pacientes, 16 con DG y 16 sin DG, que hubieran acudido a ambas visitas (basal y parto) y se parearon por edad, edad gestacional y IMC previo.

También se han incluido en el estudio las participantes de la cohorte EPIPREG, como grupo de validación y para el cálculo del MRS (ver más adelante).

Procedimiento y variables cohorte EPIDG

Las pacientes fueron derivadas a la unidad de Diabetes y Embarazo, tras un test de O'Sullivan positivo, donde se les realiza el diagnóstico de diabetes gestacional mediante SOG-100gr entre las semanas 24-30 de embarazo. Se siguieron los criterios diagnósticos de DG según el NDDG (tabla 1-introducción). Se pautan distintas visitas a lo largo del estudio:

1. Visita inicial (semana 24-30 de embarazo): Durante la realización de la SOG 100gr, las pacientes son informadas, y en caso de aceptación, firman consentimiento para participar en el estudio. A las pacientes que aceptan la participación, se les realiza lo siguiente:

- a. Evaluación general: Antecedentes familiares de DM2, DM1 o DG en familiares de primer grado. Antecedentes personales y obstétricos. Exploración clínica y antropométrica estandarizada. Información sobre el estilo de vida: consumo de tabaco, alcohol y fármacos. Cuestionario de adherencia a dieta mediterránea y de actividad física.

Toma de muestra de sangre para aislamiento de suero y células periféricas. Se guardan a -80°C alícuotas de dichas muestras para la formación de una seroteca y ADNteca. Se determinan parámetros bioquímicos, lipídicos y de inflamación. La resistencia a la insulina se mide mediante el índice HOMA (113).

- a. Una vez conocido el resultado de la SOG-100gr las pacientes son clasificadas de la siguiente forma:
 - i. Cohorte expuesta: aquellas que cumplen criterio diagnóstico para DG según NDDG. Estas pacientes harán el seguimiento habitual y rutinario de la Unidad de Diabetes y Embarazo. A lo largo de este trabajo este grupo será conocido como DG.
 - ii. Cohorte no expuesta: aquellas que no cumplen los criterios de diagnóstico para DG tras SOG-100gr. Estas pacientes serán citadas en las siguientes visitas.
- b. Variables recogidas en esta visita: peso previo al embarazo, pero actual, talla, tensión arterial, gestaciones previas, DG en gestaciones previas, RN macrosómicos en gestaciones previas, antecedente familiar de primer grado con DM2, tabaco, alcohol, enfermedades concomitantes,

tratamientos concomitantes (incluidos suplementos vitamínicos o de hierro). En el grupo DG se entrega glucómetro que cumpla la normativa ISO-2013 para mayor fiabilidad y exactitud de los datos.

2. Visita preparto (semana 36-38 de embarazo): en esta visita se vuelve a realizar la misma exploración física que en la primera visita y se obtiene de nuevo la información clínica, antropométrica y de estilos de vida. Se repite la extracción analítica. Se les solicita a todas las pacientes una SOG (75g) para realizar a partir de las 8-10 semanas tras el parto.
 - a. Variables clínicas recogidas en esta visita: además de las mismas que en la visita basal, se recogen enfermedades concomitantes de nueva aparición, acontecimientos obstétricos relevantes desde la visita anterior (ej: amenaza de parto que requiera maduración pulmonar con corticoides). Datos biométricos de la ecografía más reciente: edad gestacional en el momento de la realización de dicha ecografía, peso fetal estimado y percentil fetal.
 - b. En la cohorte expuesta se recogen, además, datos de control glucémico mediante la descarga de datos del glucómetro y del tratamiento (dieta o insulina).
3. Visita posparto (8-10 semanas)
 - a. Se realiza una SOG-75gr para la reclasificación del estado metabólico de las pacientes con DG.
 - b. Se realiza extracción de sangre a la madre para aislamiento de suero, plasma y células periféricas.

- c. Se recogen variables obstétricas y perinatales con los informes aportados por la paciente. En caso preciso se revisa las historias de partos y neonatología.
- d. Se obtiene una muestra de saliva del niño para su posterior extracción de ácidos nucleicos.
- e. Variables clínicas recogidas:
 - i. Madre: peso, tensión arterial, fecha de parto, tipo de parto (espontáneo/inducido), causa de la inducción (obstétrica, metabólica, electiva), parto vaginal/cesárea, instrumental/natural. Amenazas de parto que hayan indicado corticoides para la maduración pulmonar fetal.
 - ii. Recién nacido: peso al nacer, talla al nacer, perímetro cefálico al nacer, tipo de lactancia (natural/artificial/mixta), incidencias que hayan prolongado hospitalización del niño.

4. Visitas año y dos años

- a. Se realiza una SOG-75gr para la reclasificación del estado metabólico de las pacientes y valorar evolución a DM2 o a estados prediabéticos de las pacientes que tuvieron DG.
- b. Se realiza extracción de sangre a la madre para aislamiento de suero, plasma y células periféricas.
- c. Exploración clínica y antropométrica tanto en la madre como en el niño. Los estándares para las medidas del niño son las establecidas por la OMS (114). Para facilitar la adherencia al estudio, la madre y el niño son citados a la vez, y la exploración se realiza conjuntamente en la misma visita.

- d. Se recogió información sobre tipo de alimentación del niño y de la madre.
- e. Se obtuvo muestra de saliva del niño para su posterior extracción de ácidos nucleicos.
- f. Variables clínicas recogidas en esta visita:
 - i. Madre: peso, tensión arterial, cuestionario de actividad física y dieta mediterránea.
 - ii. Niño: tipo de lactancia, peso, talla, perímetro cefálico.

Cohorte EPIPREG

Se seleccionó un subgrupo de 480 mujeres de la cohorte de embarazadas de STORK Groruddalen (STORK G). STORK G incluyó a 823 mujeres sanas de diferentes orígenes étnicos (europeas, sudasiáticas, africanas, de Oriente Medio y sudamericanas) que acudieron a tres clínicas de salud infantil diferentes en el área de Groruddalen, Oslo, Noruega, durante el periodo 2008-2010 (115). El origen étnico se definió por el país de nacimiento del individuo o el país de nacimiento de su madre si ésta había nacido fuera de Europa (18). La cohorte EPIPREG está formada por aquellas participantes que disponían de datos de glucosa en ayunas y a las que se les había realizado un análisis de genotipado. Tiene un total de 312 mujeres europeas (EPIPREG_EU), de las cuales 73 fueron diagnosticadas de DG, y 168 sudasiáticas (EPIPREG_SA), de las cuales 68 fueron diagnosticadas de DG. Se recogieron muestras de sangre en ayunas y se realizó a todas las mujeres una prueba de tolerancia oral a la glucosa (SOG) de 75 g (prueba universal) en la semana 28±2 del embarazo. Para el presente estudio, la DG

se clasificó según los criterios ligeramente modificados de la Asociación Internacional de Diabetes y Embarazo (IADPSG) (glucosa en ayunas $\geq 5,1$ mmol/l y/o glucosa a las 2 horas $\geq 8,5$ mmol/l, ya que no se disponía de valores de glucosa a la hora).

El Comité Regional Noruego de Ética de la Investigación Médica Sanitaria del Sureste aprobó el estudio STORK-G incluyendo datos genéticos y epigenéticos (número de ref. 2015/1035). Se obtuvo el consentimiento informado por escrito de todos los participantes antes de cualquier procedimiento relacionado con el estudio.

Procesamiento de las muestras y determinación de las variables de estudio

Procesamiento de muestras:

Se extrajeron muestras de sangre periférica de la que se hicieron distintas alícuotas:

- 1- Sangre
- 2- Plasma.
- 3- Suero
- 4- Células de la línea blanca.

El plasma se obtuvo por centrifugación a 4000 xg a 4°C durante 10 minutos. Las células se obtuvieron por centrifugación a 1800 xg durante 30 minutos sin aceleración ni frenado, posteriormente se realizaba dos centrifugaciones de lavado con suero fisiológico a 1400xg durante 10 minutos cada una, con aceleración y frenado. Todas las muestras se almacenaron en una seroteca creada para este estudio a -80°C.

Determinaciones bioquímicas

La glucosa se midió usando el kit de la glucosa-oxidasa de Bayer, Leverkusen, Alemania. La insulina se midió por un radioinmunoanálisis de BioSource International Inc. (Camarillo, CA, EE.UU.). Se calculó el índice HOMA-IR (homeostasis model assessment of insulin resistance index, en inglés) según lo descrito por Matthews et al (20). La 25(OH)D se determinó mediante un kit de inmunoensayo enzimático (ELISA) de Immundiagnostik, Bensheim, Alemania. El colesterol, los triglicéridos y las HDL se midieron mediante kits enzimáticos comerciales.

Determinación de variables clínicas y antropométricas

Las variables clínicas y antropométricas que se recogen son las siguientes:

- Clínicas: edad, sexo, antecedentes familiares y personales de enfermedades metabólicas, tensión arterial. La presión arterial se midió dos veces con un esfigmomanómetro con un intervalo de 5 min entre mediciones y la media de las dos mediciones se utilizó en los análisis
- Antropométricas:
 - o Madre: Peso (anterior al embarazo y en cada visita), talla, IMC: Kg/m².
 - o Descendencia: Tipo de lactancia (exclusiva, mixta o artificial), talla, peso, perímetro cefálico, y sus respectivos percentiles.
- Nutricionales y hábitos: tabaco, alcohol, fármacos, adherencia dieta mediterránea, ejercicio físico.

Aislamiento de ADN y bisulfitación

- 1) Cohorte EPIDG. La extracción de ADN se realizó con el kit Quiamp DNA Blood Mini Kit (Qiagen, Hilden, Alemania) siguiendo las instrucciones del fabricante. En el caso de las muestras de sangre se partió de alícuotas de sangre (volumen = 200 µl) conservadas a -80°C. Las muestras de saliva se recogieron en hisopos especiales con líquido de conservación (Real Swabs Sample Collection, RBMEG) y se conservaron a temperatura ambiente hasta su posterior extracción de ADN. La calidad y la concentración del ADN se midieron utilizando el fluorómetro Qubit 3.0 con el kit Qubit dsDNA HS Assay (Thermo Fisher Scientific, Waltham, MA, EE.UU.). Un total de 500 ng de ADN genómico se trató con bisulfito con el kit Zymo EZ-96 DNA Methylation™ (Zymo Research Corp, Irvine, CA, USA) para el posterior análisis de metilación del ADN, según instrucciones del fabricante.
- 2) Cohorte EPIPREG. Las muestras de sangre se alicuotaron y almacenaron en biobanco a -80°C. El ADN de los leucocitos de sangre periférica se extrajo posteriormente, en el Laboratorio de Hormonas del Hospital Universitario de Oslo y se almacenó a -80°C. Se utilizó el kit EZ DNA methylation TM (ZYMO Research, Tustin, CA, EE.UU.) para la conversión con bisulfito del ADN.

Análisis de la metilación del ADN

La metilación del ADN en ambas cohortes se cuantificó *con Infinium Methylation EPIC BeadChip* (Illumina, San Diego, EE.UU.). Este array permite cuantificar más de 850.000 posiciones cubriendo promotores, y

regiones codificadoras de genes (83). Por su cobertura, se considera de elección de los estudios de EWAS.

Se utilizó el software estadístico R (<https://www.rstudio.com/>) para extraer los datos, utilizando el paquete ChAMP versión 2.9.10 (94) y el paquete de Bioconductor minfi (116). Las sondas se eliminaron utilizando el siguiente umbral: un p-valor de detección superior a 0,01 en una o más muestras, con un beadcount inferior a 3 en al menos el 5% de las muestras, sondas no-cpg y sondas con SNPs (89). La normalización intracelular se realizó mediante el método de normalización de cuantiles de mezcla beta (BMIQ). Se utilizó la corrección de Houseman para corregir las diferencias en metilación resultantes de diferencias en la heterogeneidad celular (117). Por último, para evitar posibles falsos positivos al encontrar DMPs se calculó la inflación del array con el paquete RaMWAs en el estudio R. Este paquete nos devuelve el valor lambda, y se consideran los siguientes rangos:

- λ entre 0.98-1.10: no hay inflación
- λ entre 1.10-1.19: inflación ligera
- λ entre 1.20 a 1.40: inflación moderada
- $\lambda > 1.4$: inflación alta

Los arrays analizados y el número de muestras se muestran a continuación:

Array Sangre materna

	Tipo de muestra	Visita basal (semana 24-48)	Visita Preparto (semana 37-38)
Embarazada DG	Sangre	16	16
Embarazada no DG		16	16

Array saliva descendencia

	Tipo de muestra	1ª visita post parto (8 semanas)	1 año pos parto
Descendencia de madres DG	Saliva	15	15
Descendencia de madres no DG		15	15

Array sangre materna cohorte EPIPREG

EPIPREG	Tipo de muestra	Europeas (Semana 24-28)	Sudasiáticas (Semana 24-28)
Embarazada DG	Sangre	73	68
Embarazada no DG		239	100

Análisis estadístico

Para todo el análisis estadístico se utilizó el programa informático R (4.0.4).

Estadística descriptiva de las cohortes

Dependiendo del tamaño de la muestra y de la distribución de normalidad de las variables se realizaron pruebas paramétricas o no paramétricas. Para saber si las variables seguían una distribución normal se utilizó la prueba Shapiro-Wilk. En el caso de las variables continuas y paramétricas se utilizó la prueba ANOVA, mientras que para las variables categóricas se utilizó la prueba chi cuadrado. Para las variables continuas no paramétricas se utilizó U de Mann Whitney.

Obtención β /M-values y extracción de los DMPs

El nivel de metilación de un sitio CpG se puede expresar en forma de valores beta o M. Los β se calculan mediante la fórmula que se expresa a continuación y oscilan entre 0 y 1. Los valores M se calculan mediante la transformación logarítmica de los β -value. Los β -value se usan generalmente para las representaciones gráficas, mientras los M-values

suelen usarse para los análisis estadísticos, aunque no existe un consenso sobre esto y ambos son válidos.

$$\beta - value = \frac{M}{M + U + \alpha} \quad M - value = \log_2\left(\frac{M}{U}\right)$$

$M = \text{Intensidad de metilación}$
 $U = \text{Intensidad desmetilada}$
 $\alpha = \text{Constante de compensación (default 100)}$

Figura 8- Fórmula para obtener los M-values a partir de los Beta values obtenidos en el array.

Las DMPs se obtuvieron utilizando el paquete limma en el estudio R (4.0.4). Para identificar los DMPs que se mantenían en el tiempo se construyó un diagrama de Venn utilizando la herramienta Venny 2.1 (<https://bioinfogp.cnb.csic.es/tools/venny/>).

Análisis de los datos obtenidos del array de metilación en sangre materna en EPIDG y EPIPREG

Los DMPs se filtraron por False Discovery Ratio (FDR) <0,05, la diferencia de media del β -value entre grupos ($\Delta\beta$) superior al 5% en términos absolutos, y el log odds de que el sitio CpG esté deferencialmente metilado (B) superior a 0. Los DMPs se obtuvieron ajustando por variables previamente asociadas a la DG como edad, IMC pregestacional, sexo fetal, ganancia de peso, edad gestacional y HOMAIR (118).

Análisis de los datos obtenidos del array de metilación en saliva de la descendencia EPIDG

El umbral de significación para determinar los DMP se fijó en un valor p inferior a 0,01 y $\Delta\beta \geq 5\%$. Se obtuvieron los DMPs de ambos tiempos (visita posparto y al año), ajustando por el sexo del bebé, tratamiento de la madre durante el embarazo (dieta o insulina) y ganancia de peso de la madre.

Estudio de correlaciones

Para poder conocer si los DMPs o CpGs de interés podían tener relación con las variables bioquímicas y/o antropométricas, se realizaron análisis de correlación. Dependiendo de la normalidad de los datos se utilizó la prueba de Spearman o de Pearson (no paramétrica y paramétrica respectivamente). La matriz de correlación se obtuvo y representó con el paquete GGally en R studio (4.0.4), y el umbral fue $r > 0.5$ y $p < 0.05$.

Curvas ROC y regresión logística

En el estudio llevado a cabo para identificar potenciales biomarcadores de DG a lo largo del embarazo se calculó la capacidad discriminativa de los CpGs y variables estudiadas. Para ello se realizaron curvas *Receiver Operating Characteristics* (ROC) para determinar el AUC de los CpGs metilados diferencialmente entre gestantes con DG y sin DG. Se seleccionaron las variables con $AUC \geq 0,8$ y $p < 0,05$. Para generar los modelos predictivos se utilizó la regresión logística binomial y como variable dependiente la DG. El análisis ROC ha permitido evaluar los modelos generados utilizando diferentes métricas como sensibilidad, especificidad, precisión y área bajo la curva (AUC).

Puntuación de riesgo de metilación (MRS)

El MRS se define como la suma ponderada de los β -value de los DMPs individuales de un número preseleccionado de sitios CpGs (ver introducción). Para construir el MRS usaron los valores ponderados de los sitios CpGs descubiertos en un meta-análisis de EWAS para la DM2 en cinco cohortes europeas prospectivas ($n=5.859$) (119). Los coeficientes de regresión para el EWAS de la DM2 se basaron en los β -value. Calculamos el MRS para cada uno de los sitios CpG descubiertos disponibles en nuestras

dos cohortes. Debido a los filtros de control de calidad de los CpG en cada cohorte, sólo 42 de los 72 CpGs pudieron ser incluidos en el MRS (Tabla Suplementaria 1). El MRS se construyó multiplicando el coeficiente de regresión del EWAS de la DM2 por los valores β individuales en EPIPREG y EPIDG (EPIPREG-UE, EPIPREG-SA y EPIDG). Después, resumimos la puntuación individual para cada uno de los 42 sitios CpGs para obtener el MRS.

Se utilizó un modelo logístico mixto de regresión para dilucidar si el MRS se asociaba con la DG en nuestros tres grupos, EPIPREG-UE, EPIPREG-SA y EPIDG. Se realizaron regresiones de modelos lineales mixtos para probar la asociación entre el MRS y variables continuas como HOMA-IR, péptido C, glucosa en ayunas y nivel de glucosa a las 2 h. Utilizamos los paquetes R lme4 y lmerTest (120) para realizar la regresión de modelos mixtos. En estos modelos, la ascendencia (europea, sudasiática, mediterránea (EPIDG)) se trató como un intercepto aleatorio para ajustar las posibles diferencias relacionadas con la ascendencia en la metilación del ADN.

Para identificar los sitios CpG individuales asociados con la DG en las poblaciones estudiadas, realizamos regresiones logísticas para cada CpG utilizado en el MRS en los tres grupos, seguido de un metaanálisis de los tres grupos para cada sitio CpG utilizando METAL.

Análisis funcional

Se realizaron análisis funcionales para identificar en qué rutas biológicas o mecanismos podrían estar implicado los genes correspondientes a los DMPs. Para ello se buscó la realizó una ontología de genes (GO) y la enciclopedia de Kyoto de genes y genomas (KEGG). Los términos GO y las

vías KEGG se obtuvieron utilizando el paquete Missmethy1 en el software R. Otra herramienta que también se ha utilizado a lo largo del estudio ha sido ENRICH (102), esta herramienta permite hacer un análisis de enriquecimiento mediante una interfaz web, con la que se resume la función en un conjunto de genes. Además, se realizó una red en STRING (<https://string-db.org/>) para identificar las interacciones proteína-proteína (PPI). El umbral significativo fue una puntuación de interacción $>0,7$ y $FDR < 0,05$. Para encontrar clusters de interés se utilizó Cytoscape (<https://cytoscape.org/>) con la aplicación MCODE. Finalmente, cada cluster fue analizado por DAVID (<https://david.ncifcrf.gov/summary.jsp>) para identificar la función potencial.

mQTLs

Para ver si nuestros sitios de metilación pueden estar influidos por variaciones en la secuencia génica, realizamos búsquedas en GoDMC de los CpGs significativos obtenidos en el metaanálisis para identificar loci de rasgos cuantitativos de metilación (mQTL). GoDMC proporciona una lista de SNPs que han sido asociados con el CpG de interés. Utilizamos la herramienta LDlink-SNPclip (121) para identificar SNPs en desequilibrio de ligamiento ($r^2 > 0,8$, $MAF = 0,01$) y las poblaciones seleccionadas fueron europeas y sudasiáticas. Realizamos búsquedas en Phenoscanner (122) de fenotipos nominalmente asociados ($P.val < 0,05$) con el mQTL o SNPs en alto desequilibrio de ligamiento. Phenoscanner es una base de datos que contiene resultados públicamente disponibles de GWAS a gran escala.

Resultados

Objetivo 1- Estudio de los patrones de metilación del ADN en muestras de sangre periférica de mujeres gestantes con y sin DG, tanto en el momento de diagnóstico de DG como al final del embarazo

A continuación, se muestran los resultados de este objetivo que han dado lugar a dos publicaciones. El diseño para el estudio de metilación del ADN mediante array de Illumina fue el siguiente:

	Tipo de muestra	2º Trimestre (semana 24-28)	Preparto (semana 37-38)
Embarazada DG	Sangre	16	16
Embarazada no DG		16	16

De todas las participantes se seleccionaron 32 pacientes, 16 con DG y 16 sin DG, que hubieran acudido a ambas visitas (basal y parto) y se parearon por edad, edad gestacional y IMC previo.

1.1- Identificación de marcas epigenéticas involucradas en rutas relacionadas con DG.

Resultados publicados en: Epigenome wide association study in peripheral blood of pregnant women identifies potential metabolic pathways related to gestational diabetes.

The aim of this objective is to identify epigenetic marks involved in the mechanisms or pathways related to gestational diabetes. A total of 32 pregnant women were selected, 16 of them with DG and 16 non-DG. DNA methylation pattern was obtained from Illumina Methylation Epic BeadChip, from peripheral blood samples at the diagnostic visit (26-28 weeks). Differential methylated positions (DMPs) were extracted using ChAMP and limma package in R 2.9.10, with a threshold of $FDR < 0.05$, $\Delta\beta > |5\%|$ and $B > 0$. A total of 1.141 DMPs were found, and 714 were annotated in genes. A functional analysis was performed, and we found 23 genes significantly related to carbohydrate metabolism. Finally, a total of 27 DMPs were correlated with biochemical variables such as glucose levels at different points of oral glucose tolerance test, fasting glucose, cholesterol, HOMAIR and HbA1c, at different visits during pregnancy and postpartum. Our results show that there is a differentiated methylation pattern between DG and non-DG. Furthermore, the genes annotated to the DMPs could be implicated in the development of DG as well as in alterations in related metabolic variables.

Característica de la cohorte

Las características de los sujetos se representan en la Tabla 2. Ambos grupos estaban pareados por edad, edad gestacional e IMC previo al embarazo. El grupo control mostró un mayor aumento de peso durante el embarazo en comparación con el grupo con DG. El resto de las variables no fueron estadísticamente diferentes entre los grupos. En el grupo de DG, tras la recomendación de dieta y ejercicio, 6 gestantes precisaron tratamiento farmacológico adicional (insulina).

Análisis de metilación diferencial

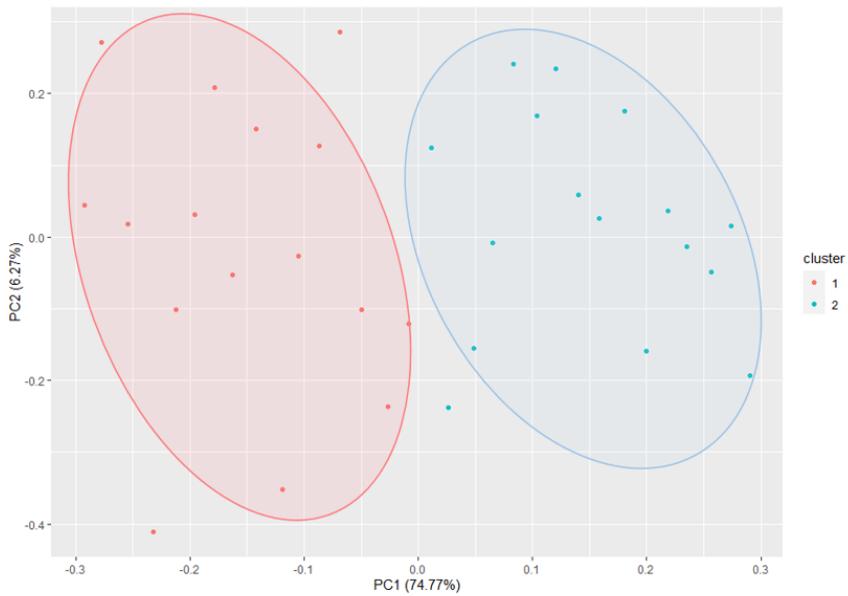
Se identificaron un total de 1.141 DMPs al comparar DG frente no-DG, con un umbral de $FDR < 0.05$, $\Delta\beta > |0.5|$ y $B \geq 0$. Se observó un ligero nivel de inflación en nuestro EWAS ($\lambda = 1.136$) (Figura suplementaria 1). El análisis de PCA, muestra claramente como se separan ambos grupos (Figura 9A). En el volcano podemos apreciar los CpGs más diferenciados según estén hiper o hipometilados (figura 9B), observándose que los cg13824270 (anotado en el gen factor de procesamiento de pre-ARNm 4B, PRPF4B), cg20409752 (anotado en CAMTA1) y cg050373962 son los más hipometilados; mientras que los cg04248279 (anotado en el gen similar a rabfilina, 3A RPH3AL), cg20758759 (en gen no anotado) y cg02518222 (anotado en gen hidroxisteroide 17-beta deshidrogenasa 11, HSD17B11) son los más hipermetilados (Tabla suplementaria 2).

Tabla 2- Características de la cohorte en la visita basal, parto y posparto

Comparación	Basal (24/28 semanas)		Pre-prarto (36-38 semanas)			Pos-parto (8 semanas)			
	Non-GDM	GDM	P-value	Non-GDM	GDM	P-value	Non-GDM	GDM	P-value
Edad	34.2±4.5	33.8±4.1	0.808						
Edad gestacional (semanas)	29±3	28±2	0.579	36±2	36±1	0.282			
O'sullivan mmol/l (mg/dl)	8.92±0.9 (160.7±16.3)	9.6±1.23 (172.9±22.2)	0.085						
Peso (Kg)	75.4±11.3	76.0±13.5	0.892	78.7±11.9	76.5±13.1	0.622	71.9±12	67.3±11.5	0.335
IMC pre-embarazo	24.6±6.9	23.8±6.9	0.862						
PS (mm Hg)	104.2±9.9	111.4±15.4	0.131	106.9±12.8	109.1±12.8	0.643	104.1±7.32	108.3±8.9	0.179
PD (mm Hg)	69.1±8.2	70.1±7.8	0.710	71.2±6.5	73.6±8.8	0.393	78.4±6	75.9±5.4	0.242
Colesterol mmol/l (mg/dl)	7.04±1.25 (272±48.1)	6.72±1.32 (258.5±51.1)	0.446	7.09±1.27 (273.6±49.1)	6.6±1.4 (254.7±54)	0.308	5.55±0.4 (214.4±15.3)	5.01±0.91 (193.5±35.3)	0.09
HDL (mg/dl)	2.14±0.47 (82.7±18.2)	2.02±0.41 (77.8±15.7)	0.424	1.93±0.32 (74.6±12.3)	1.9±0.44 (73.2±16.8)	0.785	1.72±0.4 (66.3±15.3)	1.5±0.39 (58.1±15)	0.156
Tg (mg/dl) †	2.21±0.78 (194±68)	2.09±0.43 (183±38)	0.829	2.84±1.28 (249±112)	2.76±0.9 (242±79)	0.749	1.01±0.43 (89±38)	0.7±0.65 (61±57)	0.890
HbA1C (%)	5.1±0.3	5.3±0.4	0.151	5.3±0.3	5.8±1.6	0.045	5.4±0.3	5.4±0.41	0.756
HOMA-IR †	1.5±1.1	1.7±1.4	0.147	1.5±1.4	2.1±3.2	0.883	1.0±0.7	0.8±1.3	0.964

Los datos se expresan como media ± desviación estándar. Las proporciones se compararon mediante la prueba de chi-cuadrado y las variables cuantitativas se analizaron mediante la prueba t para datos no apareados o la prueba U de Mann-Whitney según su distribución normal. IMC: índice de Masa corporal. PS: presión arterial sistólica. PD: presión arterial diastólica. HDL: colesterol de lipoproteínas de alta densidad. Tg: triglicéridos.

A)



B)

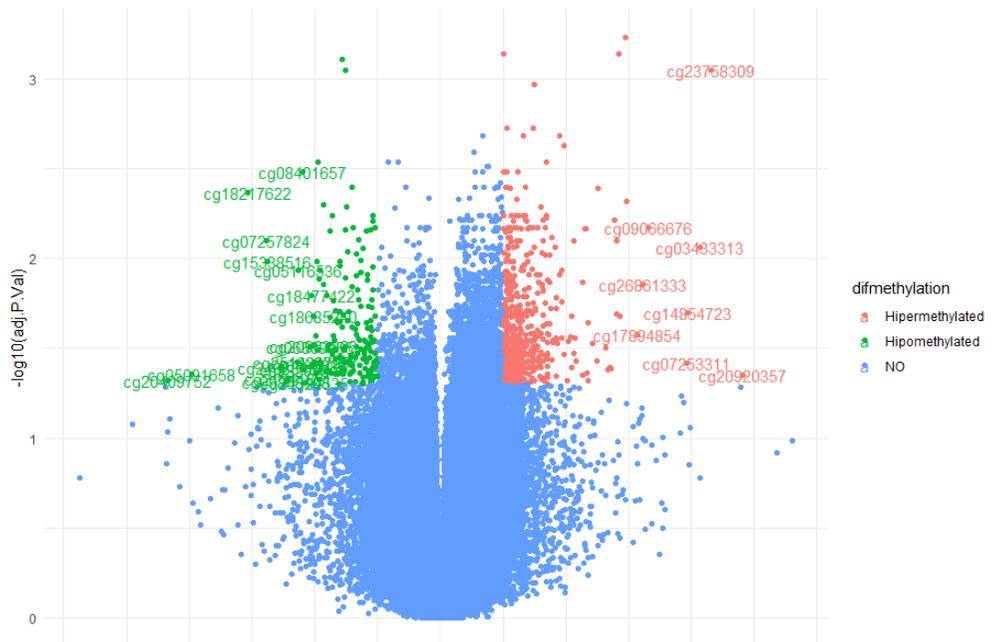


Figure 9- A) Análisis de componentes principales entre DG (azul, cluster 2) y no-DG (rosa, cluster 1). B) Gráfica volcano: verde CpGs más hipometilados significativos, rosa: CpGs más hipermetilados significativo. NO: no significativos.

De acuerdo con la ubicación de los CpGs respecto al gen (*Feature position*), el 36% se encuentra en el cuerpo del gen, el 37% en la región intergénica (IGR), alrededor del 20% en áreas promotoras (sitio de inicio de transcripción (TSS) -1500 1500, 5'UTR y TSS200) y el resto se encuentra al inicio del exón y en 3'UTR (figura 10A). En relación con la posición respecto a la isla (CGi), la mayoría de los DMPs se encontraron en la posición Open sea (67%). El resto de los DMPs se encontraron en las posiciones shore (19%), island (7%) y shelf (7%) respectivamente (figura 10B).

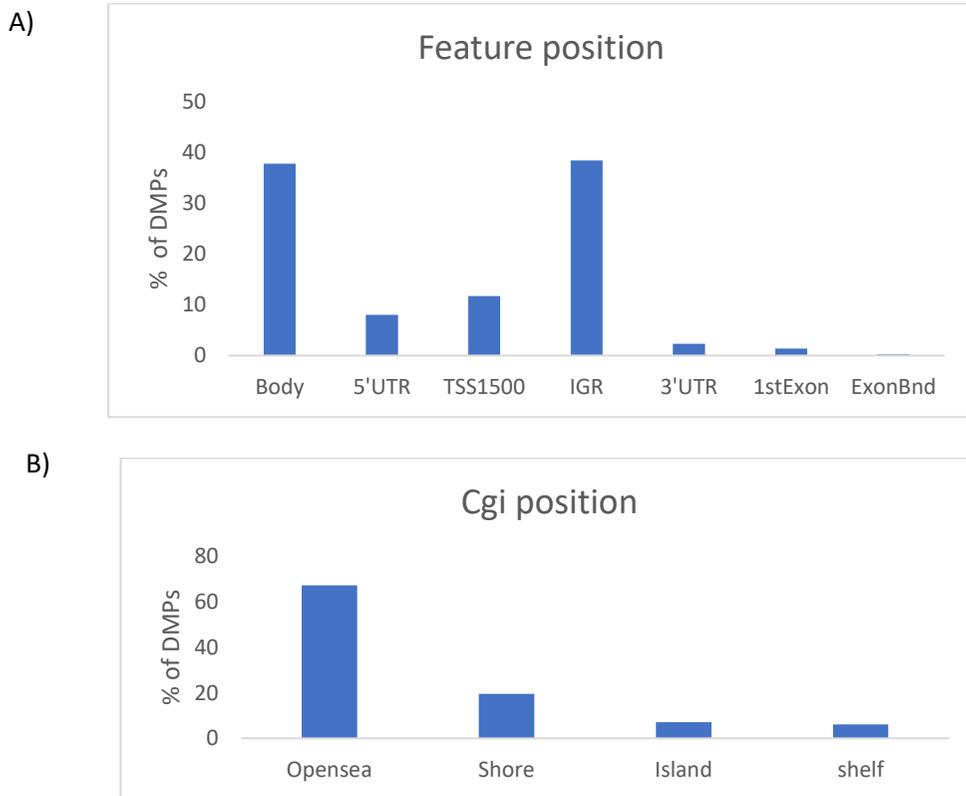


Figure 10- A) Posición de los DMPs respecto al gen (*Feature position*) B) Posición DMPs con respecto a las islas CpGs (*Cgi position*).

Análisis enriquecimiento funcional

Con el fin de conocer los posibles mecanismos involucrados en la DG, se realizó un análisis de enriquecimiento funcional. En primer lugar, se realizaron análisis de GO y KEGG en todos los DMPs con sus genes anotados, que fueron un total de 714. Los términos GO más significativos se representaron en la figura 11A. Entre los GO significativos se identificaron los siguientes mecanismos: la regulación de la secreción de insulina, la respuesta celular al estímulo de aminoácidos y la respuesta a los niveles de nutrientes. El análisis de KEGG mostró que las vías más significativas fueron la vía de señalización de la insulina, la síntesis de la hormona de crecimiento, la vía de señalización de ErbB y la diabetes tipo 2 (figura 11B).

Además, se realizó una red String en los 714 genes anotados para identificar interacciones proteína-proteína (PPI) y sus mecanismos relacionados. La red de interacciones se clusterizó según funcionalidad en Cytoscape utilizando la aplicación MCODE y se analizó en DAVID. Como se muestra en la Figura 2C, se definieron un total de 15 grupos utilizando los valores por defecto de la aplicación MCODE. Los tres grupos más significativos (valor de enriquecimiento $p < 1e-16$) estaban relacionados con vías del metabolismo de los carbohidratos: regulación positiva del proceso glucolítico, vía de señalización MAPK, vía de la insulina, respuesta celular a los niveles de nutrientes, diabetes tipo 1, diabetes tipo 2 y resistencia a la insulina. Un total de 23 de nuestros genes anotados fueron incluidos en estos grupos y también hubo interacción entre ellos.

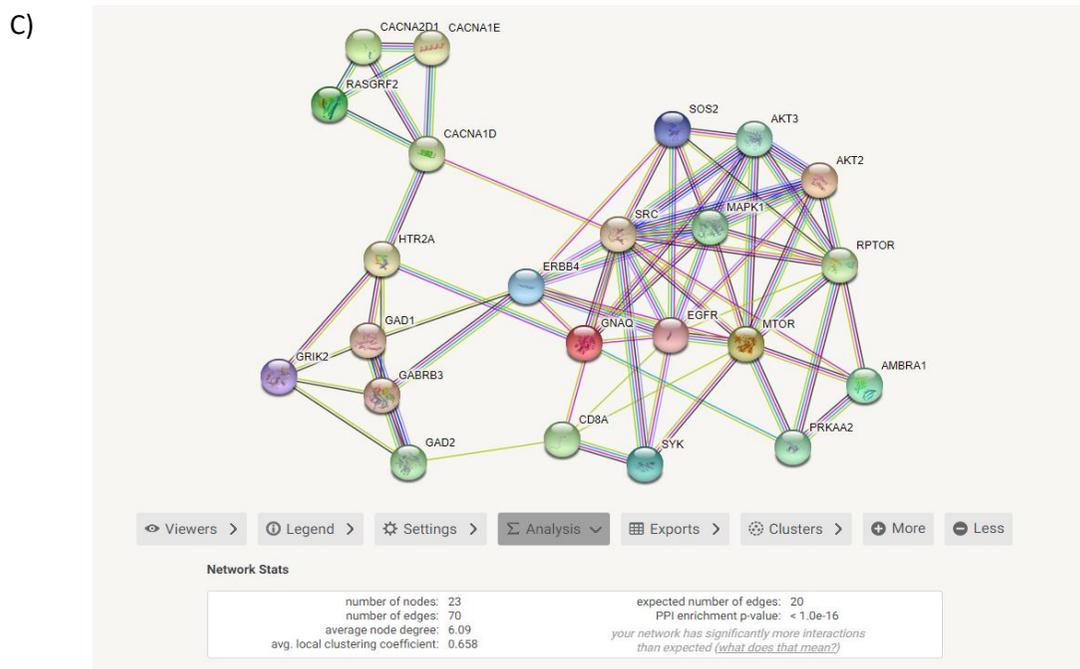
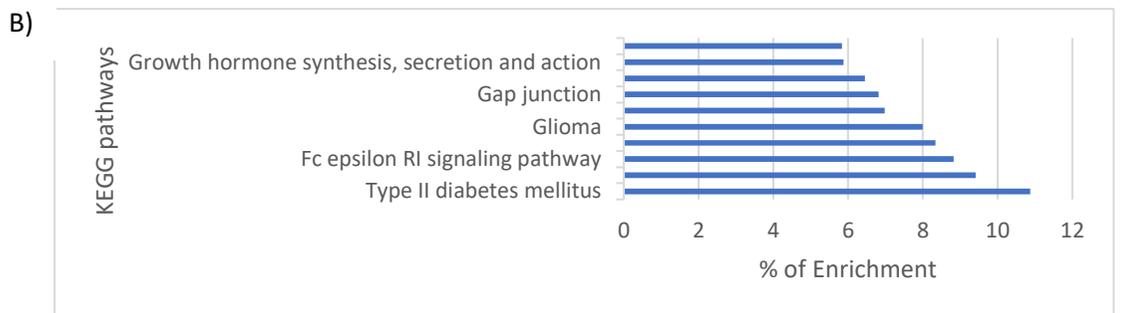
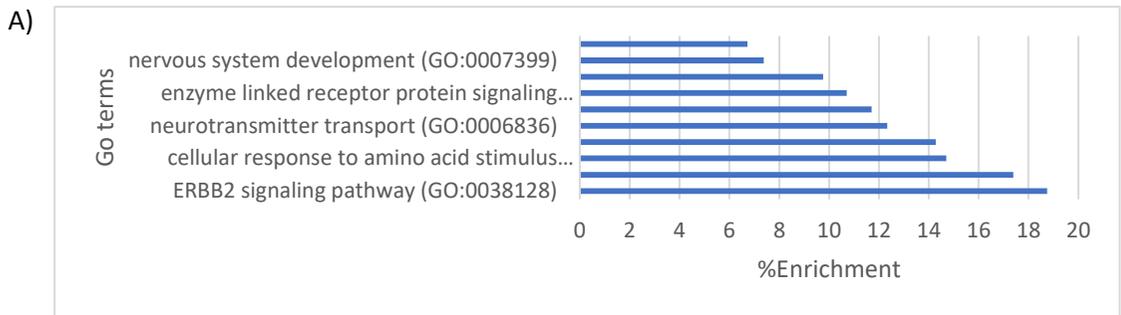


Figure 11- A) Términos GO más enriquecidos $p.val < 0.005$ B) Rutas de señalización KEGG más enriquecidos, $p.val < 0.005$. C) Red de PPI de String net PPI ($p-value < 1.0e-16$) con 23 genes relacionados con el metabolismo de la glucosa.

Análisis de correlación

A partir de los 100 CpGs más significativos, se seleccionaron genes con más de un CpG y aquellos genes relacionados con el metabolismo de los carbohidratos, según nuestro análisis de enriquecimiento, para realizar un análisis de correlación con las variables bioquímicas en diferentes momentos T0 (basal), T1 (preparto) y T2 (posparto). Un total de 27 CpGs se correlacionaron significativamente con las variables bioquímicas ($p < 0.05$, $r > |0.5|$) (Tabla 3).

Table 3- CpGs correlacionados con las variables bioquímicas.

CpG	Genes	Variables	r	P.val
cg13487126	AKT2	SOG120.T2	0.5939	0.03234932
cg00856432	CACNA1D	VitaminD.T2	0.5914	0.00072894
cg20409752	CAMTA1	HOMAIR.T0	-0.5521	0.00105251
cg17816038	CAMTA1	Try.T1	-0.5287	0.00186484
cg14688342	EGFR	SOG0.T2	-0.6216	0.02334455
cg12854232	ERBB4	Glucose.T0	-0.5263	0.00197518
cg04217177	ESRRB	SOG0.T2	0.6725	0.01178094
cg17434154	ESRRG	SOG0.T2	0.5535	0.04973052
cg07420274	GAD1	HOMAIR.T0	-0.5227	0.00214649
cg20230784	GNAQ	SOG180.T0	0.5602	0.00085423
cg20344956	LOC646214	HbA1c.T2	0.53	0.00310505
cg07894331	MAPK1	BMI_p	0.5187	0.00235336
cg10211745	MINOS1-NBL1	SOG120.T2	0.6742	0.01149194
cg19234140	MINOS1-NBL1	SOG120.T2	0.6898	0.00907725
cg09418268	MTOR	Glucose.0	0.5391	0.00145423
cg14579430	NBL1	SOG120.T2	0.6161	0.02493686
cg15589641	NBL1	SOG120.T2	0.6261	0.02205904
cg12474394	NBL1	SOG120.T2	0.6912	0.00888177
cg14715327	OR4A16	SOG120.T2	0.6704	0.01215022
cg12080079	PAX7	Chl.T1	0.5809	0.00048966
cg21873971	PDIA5	SOG0.T2	0.7919	0.00125394
cg05894391	PRKAA2	SOG0.T2	0.6177	0.02446442

cg11055991	PTPRN2	SOG0.T2	0.5803	0.03758363
cg14479198	RASGRF2	SOG0.T2	-0.5752	0.03970591
cg19891476	SYK	SOG0.T2	0.5576	0.04771194

r: correlación de Spearman. SOG00: Sobre carga oral de glucosa en ayunas. SOG120: Sobre carga oral de glucosa punto 120 min. SOG180: Sobre carga oral de glucosa punto 180 min Tg: triacilgliceridos BMI_p: índice de masa corporal pre-embarazo. Chl: Nivel de colestesrol .0: visita inicial .1: visita preparto. .2: visita pos-parto.

Las correlaciones más altas se encontraron para los valores de glucosa basales de la curva de SOG en el posparto (SOG0.2) en los genes del receptor del factor de crecimiento epidérmico (EGFR) (cg14688342), familia de isomerasa de disulfuro de proteínas A miembro 5 (PDIA5) (cg21873971), subunidad alfa 2 de la quinasa de proteínas AMP-activada (PRKAA2) (cg05894391), y en la región no anotada cg04217177 (Figura 12); y para los valores de glucosa a las 2 horas de la sobrecarga en el posparto (SOG120.2) en el gen antagonista BMP de la familia DAN (NBL1) (cg14579430, cg15589641, cg12474394, cg14715327) y en el sistema organizador de la membrana interna mitocondrial 1-NBL1 (MINOS1-NBL1) (cg10211745, cg19234140) (Figura 13).

Sobre carga oral de glucosa en ayuna en el posparto (SOG0.2)

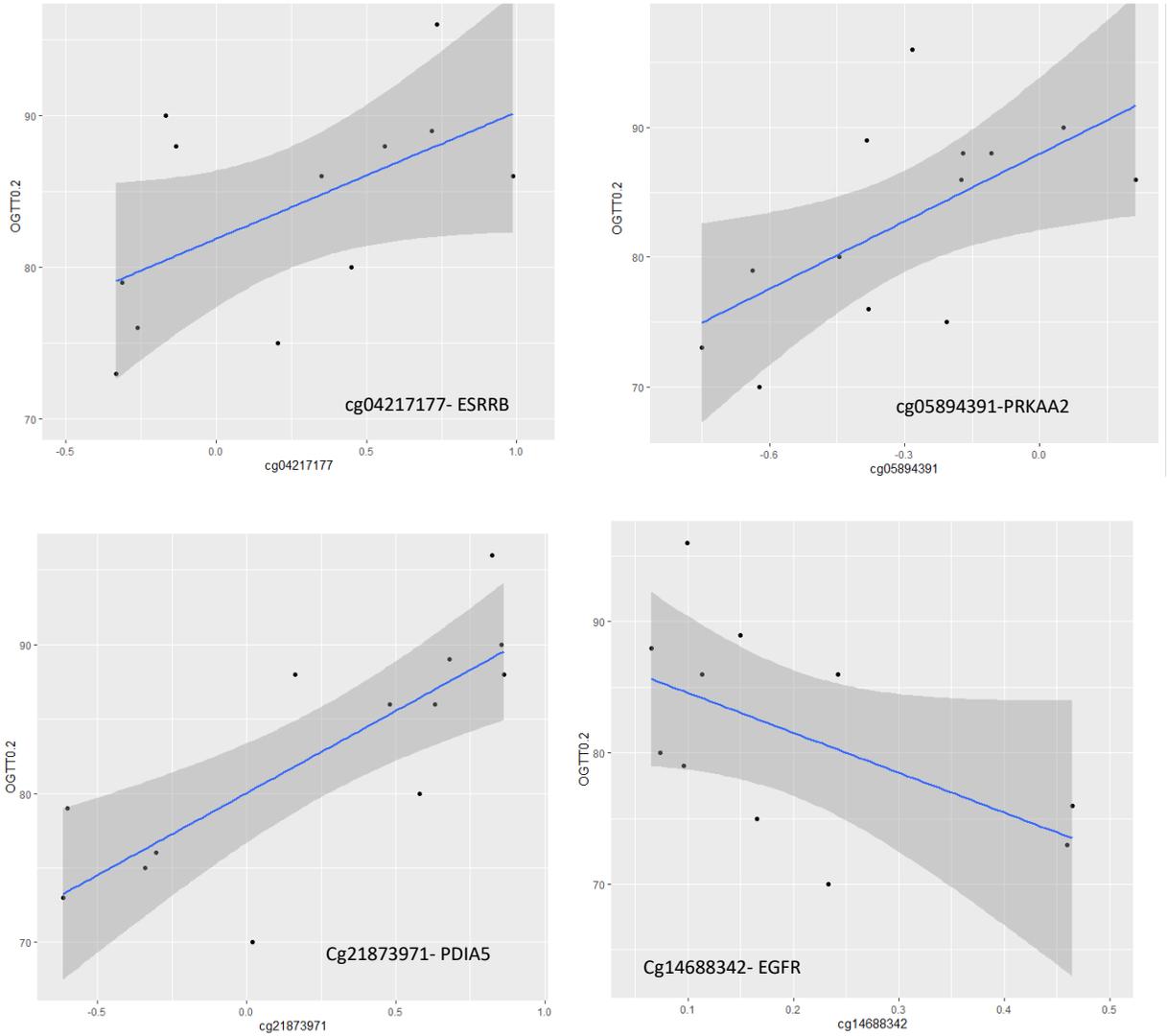


Figure 12- Gráficas de correlación de los CpGs más asociados con la sobrecarga oral de glucosa en ayunas en la visita preparto.

Sobre carga oral de glucosa 2 horas en el posparto (SOG120.2)

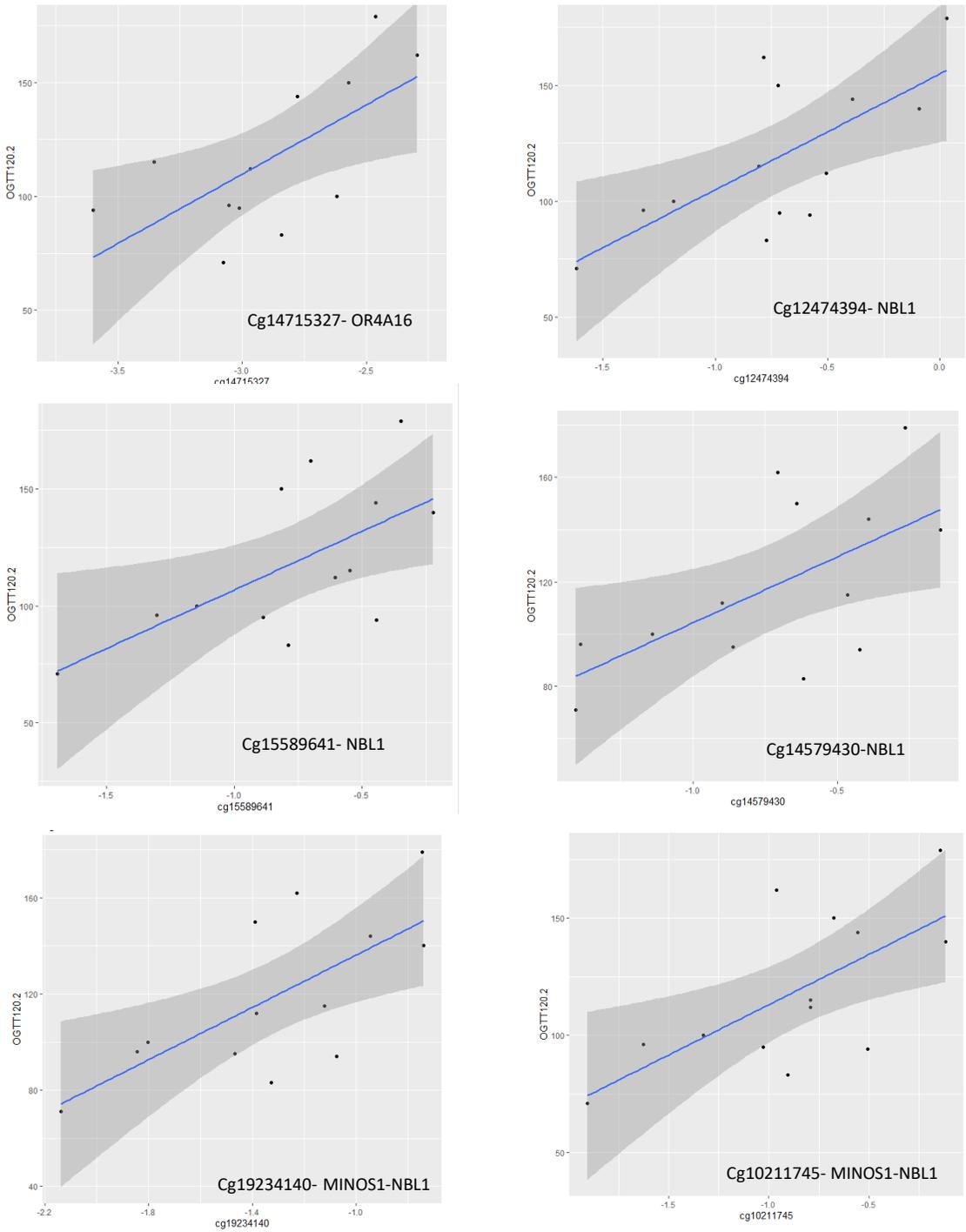


Figure 13- CpGs más correlacionados con la sobrecarga oral de glucosa en el punto 120 de la visita preparto.

1.2- Marcas epigenéticas asociadas con la DG en dos momentos durante el embarazo.

Resultados publicados en el artículo: Epigenetic marks associated with gestational diabetes mellitus across two time points during pregnancy.

In this objective, we explored the epigenetic profile in maternal peripheral blood samples through pregnancy to find potential epigenetic biomarkers for gestational diabetes mellitus (DG). We performed an epigenome-wide association study in maternal peripheral blood samples in 32 pregnant women (16 with DG and 16 non-DG) at pregnancy week 24-28 and 36-38. Biochemical, anthropometric, and obstetrical variables were collected from all the participants. The main results were validated in an independent cohort with different ethnic origin (European = 307; South Asians = 165). Two hundred and seventy-two CpGs sites remained significantly different between DG and non-DG pregnant women across two time points during pregnancy. The significant CpG sites were related to pathways associated with type I diabetes mellitus, insulin resistance and secretion. Cg01459453 (SELP gene) was the most differentiated in the DG group versus non-DG (73.6 vs. 60.9, $p = 1.06E-11$; $FDR = 7.87E-06$). Three CpG sites (cg01459453, cg15329406, and cg04095097) were able to discriminate between DG cases and controls ($AUC = 1$; $p = 1.26E-09$). Three differentially methylated positions (DMPs) were replicated in an independent cohort. To conclude, epigenetic marks during pregnancy differed between DG cases and controls along time. Three CpGs were able to discriminate DG and non-DG groups with high specificity and sensitivity, which may be potential biomarker candidates for prediction of DG in the future.

Características cohorte

La Tabla 4 presenta las características principales de la población EPIDG y EPIPREG. Como se ha visto en el apartado anterior, en EPIDG, no hubo diferencias significativas en las principales variables entre ambos grupos. No obstante, se observó que el grupo con DG experimentó una ganancia de peso menor durante el último trimestre del embarazo en comparación con el grupo de control. Esta diferencia probablemente se atribuya a un buen control metabólico y a la adherencia a las recomendaciones de estilo de vida por parte de las mujeres embarazadas con DG después del diagnóstico. Después de la implementación de las recomendaciones dietéticas, se necesitó tratamiento farmacológico adicional (insulina) en el 37.5% de los casos de DG.

Tabla 4- Características de las mujeres en las dos cohortes EPIDG y EPIPREG.

Comparación	Cohorte de descubrimiento (EPI-DG)						Cohorte de validacion (EPIPREG)								
	Basal (24/28 semanas) n=32			Parto (36-38 semanas) n=32			Complete cohort n= 472			Europeas n= 307			Sudasiáticas n= 165		
	Non-DG (n=16)	DG (n=16)	P	Non-DG (n=16)	DG (n=16)	P	Non-DG (n=331)	DG (n=141)	P	Non-DG (n=234)	DG (n=73)	P	Non-DG (n=97)	DG (n=68)	P
Edad	34.2±4.5	33.8±4.1	NS				29.4±4.6	29.4 ±4.8	NS	30.1±4.5	30.14±4.9	NS	27.9±4.5	28.7±4.7	NS
Edad Gestacional (Semanas)	27.6±2.1	28.1±2.8	NS	35.8±1.4	36.25±1.1	NS	27.9±12.9	27.9±13.0	NS	28.1 ±12.5	28.0±13.8	NS	27.5 ±10.2	27.8±12.1	NS
O'sullivan mmol/l (mg/dl)	8.92 ± 0.9 (160.7±16.3)	9.6 ± 1.23 (172.9±22.2)	NS			NS									
Glucosa 0' mmol/l (mg/dl)	4.5±0.4 (81.0 ± 7.21)	4.9±0.6 (88.28 ± 10.81)	0.05	4.3±0.8 (77.47 ± 14.41)	4.4±0.9 (79.27± 16.21)	NS	4.51±0.33 (81.25±5.95)	5.48±0.56 (98.73±10.09)	<0.001	4.48±0.32 (80.71± 5.77)	5.46±0.59 (98.37± 0.63)	<0.001	4.57±0.32 (82.33±5.77)	5.5±0.53 (99.09±9.55)	<0.001
Glucosa 60' mmol/l (mg/dl)	8.2±1.4 (147.73 ± 25.22)	11.1±1.2 (199.98± 21.62)	<0.001												
Glucosa 120' mmol/l (mg/dl)	6.9±1.4 (124.31 ± 25.22)	9.9±1.2 (178.36 ± 21.62)	<0.001				5.77±0.21 (103.95±3.78)	7.05±0.36 (127.01±6.49)	<0.001	5.7±1.1 (102.69 ±19.82)	7.04±0.9 (126.83± 17.66)	<0.001	5.94±0.45 (107.02±8.11)	7.06±0.54 (127.19±9.73)	<0.001
Glucosa 180' mmol/l (mg/dl)	5.8±1.1 (104.49 ± 19.82)	8.3±1.3 (149.53 ± 23.42)	<0.001												

Peso (Kg)	75.4±11.3	76.0±13.5	NS	78.7±11.9	76.5±13.1	NS	73.4±13.2	77.4±14.9	2.0E-03	76.5±13.2	81.7±15.3	2.0E-03	65.94±9.9	72.8±13.2	1.0E-04
IMC Previo	25.5±4.1	25.8±4.5	NS				23.8±4.32	25.6±5.37	4.4E-04	24.1±5.01	26.3±6.05	5.00E-3	23.0±3.2	24.84±4.4	6.4E-03
PS (mm Hg)	104.2±9.9	111.4±15.4	NS	106.9±12.8	109.1±12.8	NS	103.9±9.6	107.1±9.2	1.0E-03	105.8±9.5	110.1±9.2	1.0E-2	99.2±8.5	103.9±8.2	1.0E-03
PD (mm Hg)	69.1±8.2	70.1±7.8	NS	71.2±6.5	73.6±8.8	NS	67.1±7.3	68.4±6.7	3.1E-03	67.8±6.9	69.7±7.2	2.4E-2	65.48±7.6	67.1±5.8	NS
Colesterol mmol/L (mg/dl)	7.04±1.2 (271.81 ± 46.33)	6.7±1.3 (258.69 ± 50.19)	NS	7.1±1.31 (274.13±50.58)	6.6±1.4 (254.83±54.05)	NS	6.4±1.04 (247.1 ± 40.15)	5.9±1.1 (227.8± 42.47)	1.3E-06	6.5±1.1 (250.97 ±42.47)	6.01±1.1 (232.05± 42.47)	7.0E-4	6.2±1.02 (239.38±39.38)	5.7±1.0 (220.08±38.61)	8.4E-03
HDL-cholesterol mmol/L (mg/dl)	2.1±0.47 (81.08 ± 18.15)	2.01±0.4 (77.61±15.44)	NS	1.9±0.3 (73.36±11.58)	1.9±0.4 (73.36±15.44)	NS	1.9±0.4 (73.36±15.44)	1.8±0.4 (69.5±15.44)	6.8E-04	2.00 ±0.4 (77.22± 15.44)	1.8±0.4 (69.5±15.44)	1.1E-02	1.9±0.4 (73.36±15.44)	1.8±0.6 (69.5±23.17)	4.6E-02
Tg mmol/L (mg/dl)	2.2±0.6 (192.5±52.5)	2.2±0.5 (192.5±43.75)	NS	2.9±0.9 (253.75±78.75)	2.8±0.7 (245±61.25)	NS	1.9 ±0.7 (166.25±61.25)	2.0 ±0.7 (175±61.25)	NS	1.9 ±0.7 (166.25 ±61.25)	1.9±0.7 (166.25± 61.25)	NS	1.9±0.6 (166.25±52.5)	2.1±0.7 (183.75 ± 61.25)	NS
HbA1C	5.1±0.28	5.3±0.37	NS	5.3±0.3	5.8±1.6	NS	5.1±0.5	5.3±0.3	6.7E-11	5.0 ±0.3	5.2±0.3	4.3E-05	5.2±0.3	5.4 ±0.3	5.3E-05
HOMA-IR	1.6±0.7	2.1±1.2	NS	3.5±7.3	3.7±4.2	NS	1.6±0.6	2.1±0.4	1.0E-08	1.5±1.1	1.9±0.9	1.4E-02	1.7±0.7	2.3±0.5	6.0E-06
Ganancia de peso (Kg)				3.3±1.4	0.49±2.6	1.0E-03									

Los datos se expresan como media ± desviación estándar. Las proporciones se compararon mediante la prueba de chi-cuadrado y las variables cuantitativas se analizaron mediante la prueba t para datos no apareados o la prueba U de Mann-Whitney según su distribución normal. IMC: índice de masa corporal. PAS: presión arterial sistólica. PAD: presión arterial diastólica. HDL: colesterol de lipoproteínas de alta densidad. Tg: triglicéridos. El aumento de peso se calculó como la diferencia entre el peso en la visita prenatal (T1) y en la visita diagnóstica (T0). NS: no significativo

Patrón de metilación en mujeres con DG y no-DG a lo largo del embarazo

Un total de 1141 CpGs y 465 sitios CpGs (FDR <0.05, deltaBeta > 5% y B ≥ 0) mostraron diferencias significativas en la metilación del ADN entre las visitas diagnóstico (T0) y las visitas prenatales (T1), respectivamente. La mayoría (66%) de los DMPs estaban hipermetilados en ambas visitas (T0 = 757 DMPs, T1 = 311 DMPs) en el grupo DG en comparación con el grupo sin DG. Según la posición genómica, la mayoría de estos DMPs se encontraban en Open Sea y, según el contexto del gen, principalmente en el cuerpo del gen y la región intergénica (Figura suplementaria 2). Además, evaluamos si el perfil epigenético cambió durante el embarazo en ambos grupos. Observamos un total de 267 DMPs en el grupo de control a lo largo del tiempo, mientras que en el grupo DG no encontramos ningún sitio CpG diferencialmente metilado a lo largo del tiempo (Tabla suplementaria 3). Por último, observamos solo dos CpGs que respondieron de forma diferente en el tiempo en el grupo DG en comparación con el grupo sin DG (Figura suplementaria 3).

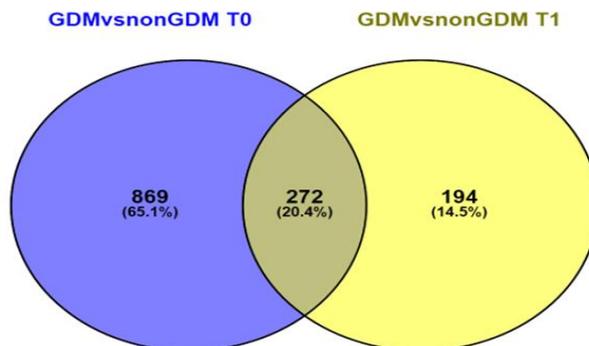


Figure 14- Diagrama de Venn con los DMPs encontrados en T0 frente a los encontrados en T1.

Observamos que 272 sitios CpGs tenían metilación diferencial entre los casos de DG y los controles en las dos visitas (Figura 14). En la tabla 5 se muestran las características de los 20 DMPs más diferenciados. Estos DMPs estaban relacionados con 12 genes (Tabla 5). Cg01459453, anotado en el gen Selectin-P (SELP), fue el más diferencialmente metilado entre ambos grupos (12.7%), estando hipermetilado en el grupo DG. El gen NBL1 estaba especialmente enriquecido con 3 sitios CpGs (cg18923740, cg15589641 y cg14579430) diferencialmente metilados (cuerpo, TSS1500 y 5'UTR). No se encontraron regiones diferencialmente metiladas (DMRs) utilizando el algoritmo del paquete Bioconductor DMRcate.

Tabla 4- 20 DMPs más diferenciados entre DG vs no-DG a lo largo del embarazo.

ID_CpG	LogFC	P.Value	FDR	Chr	Gene	feature	cgi
cg01459453	0.181	1.06E-11	7.87E-06	1	SELP	Body	opensea
cg12432693	-0.057	4.07E-11	1.51E-05	11	OTOG	Body	opensea
cg18217622	-0.206	8.73E-10	0.00016171	6		IGR	island
cg18923740	0.090	2.18E-09	0.00032293	1	NBL1	Body	island
cg12215871	0.115	2.89E-09	0.00035673	7	SDK1	Body	opensea
cg04802986	0.118	6.75E-09	0.00071486	1	LGR6	Body	opensea
cg21809624	-0.080	1.40E-08	0.00098142	17		IGR	opensea
cg08386137	0.097	1.69E-08	0.00104499	6		IGR	island
cg06279296	0.157	2.03E-08	0.00109495	10	DIP2C	Body	opensea
cg27603605	0.077	2.61E-08	0.00129099	12	TMEM132B	Body	opensea
cg15589641	0.112	3.35E-08	0.00145799	1	NBL1	TSS1500	shore
cg10102108	-0.093	3.73E-08	0.00153659	20		IGR	opensea
cg04600077	-0.082	5.04E-08	0.00169965	5		IGR	shore
cg14688342	-0.073	5.53E-08	0.00169965	7	EGFR	Body	opensea

cg07257824	-0.180	5.62E-08	0.00169965	6		IGR	island
cg12080079	-0.132	6.27E-08	0.00177299	1	PAX7	Body	opensea
cg23743013	0.078	7.00E-08	0.00177949	3		IGR	opensea
cg14579430	0.120	7.09E-08	0.00177949	1	NBL1	5'UTR	shore
cg01743873	-0.081	8.85E-08	0.0019774	11	CD151	5'UTR	shore
cg01757548	0.084	9.00E-08	0.0019774	6		IGR	island

Análisis de ontología génica (GO)

Se realizó un análisis de enriquecimiento de conjuntos de genes utilizando EnrichR y GO (100,102). Los 272 sitios CpGs se anotaron en 140 genes. De ellos, el 10% contenía más de una sonda con metilación diferencial (Figura suplementaria 4). El análisis de KEEG identificó seis vías significativas relacionadas con la Guía de Axones, la vía de señalización ErbB y la vía de señalización del calcio, entre otras. Además, se identificaron vías relevantes relacionadas con la Diabetes Mellitus tipo I (PTPRN2), resistencia a la insulina (RPS6KA2) y secreción de insulina (ADCYAP1R1) (Figura 14). El análisis basado en la base de datos de Expresión de Proteínas en Tejidos reveló que un grupo de genes anotados en nuestros DMPs, como CHSY1, ERBB4, TRIP6, DIP2C, AGRN y LGR6, mostraron niveles de expresión más altos en la placenta que en otros tejidos (Figura 15).

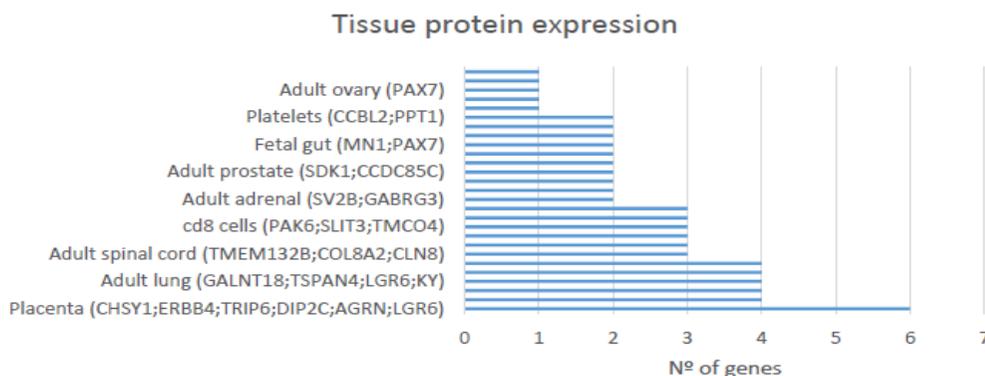
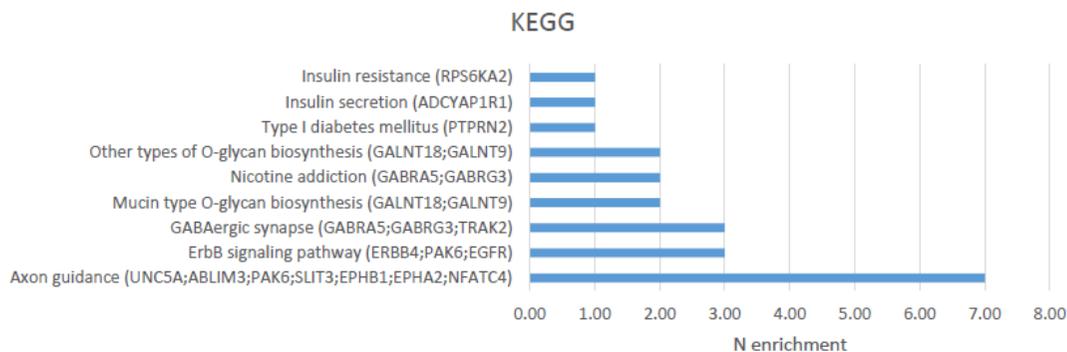


Figure 15- A) Rutas de señalización en KEGG más enriquecido ordenado por valor inverso del p -val
 B) Número de genes por expresión proteica en tejido. Ordenado por el valor inverso de p -val.

Análisis de curva ROC de DMPs

Seleccionamos los 272 sitios CpGs y realizamos un análisis de curva ROC para identificar si estos podían diferenciar entre los casos de DG y los controles sin DG. De los 272 sitios CpG, 104 pudieron discriminar estadísticamente entre los casos de DG y los controles ($AUC \geq 0.8$) en la visita diagnóstica (T0). El mismo enfoque se realizó con las siguientes variables clínicas: HOMA-IR, edad, edad gestacional, IMC pregestacional y los valores de glucosa en cada punto de la SOG. Entre las variables clínicas, solo los valores de glucosa de la SOG tuvieron un AUC estadísticamente significativo ≥ 0.8 ($AUC_{SOG_60} = 0.047$; $AUC_{SOG_120} = 0.025$; $AUC_{SOG_180} = 0.043$). Realizamos un análisis de regresión logística para

identificar las variables que mejor discriminaban entre los casos de DG y los no DG, incluyendo todos los sitios CpGs y los valores de glucosa de la SOG que tenían curvas ROC significativas. De todas las variables, 3 CpGs (cg01459453, cg15329406, cg04095097) se mostraron como los discriminadores más significativos (Tabla 6). Las figuras ROC se muestran en la Figura suplementaria 5. Las mujeres embarazadas con DG mostraron una metilación significativamente más alta en cg01459453 (SELP) en comparación con las mujeres sin DG (73.6% vs 60.9%, $p = 1.10 \cdot 10^{-7}$). La misma tendencia significativa se observó para cg15329406 y cg04095097 (Figura 16).

Tabla 6-Resultados del análisis ROC

Model	Sensitivity	Specificity	Precision	AUC	P-value
Model 1 cg01459453	93.8	93.8	93.75	0.969	3.45E-08
Model 2 cg01459453, cg15329406	100	100	100	1	2.32E-10
Model 3 cg01459453, cg15329406, cg04095097	100	100	100	1	1.26E-09

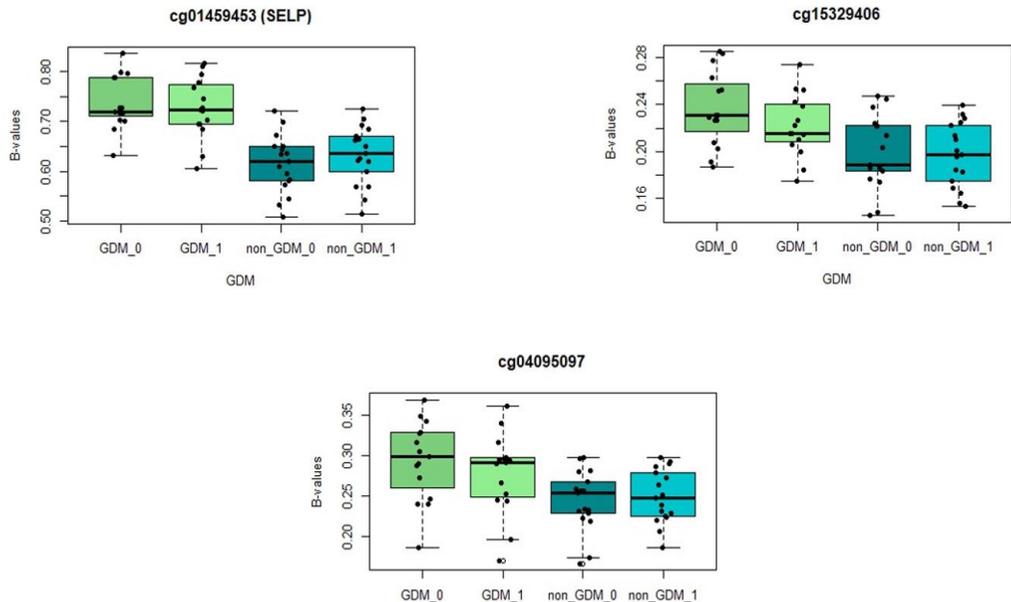


Figure 15- Gráfica Box-plot de los tres CpGs estadísticamente significativos incluidos en la regresión logística. Leyenda: DG_0: Grupo diabetes gestacional visita T0. DG_1: Grupo diabéticas gestacionales visita T1. Non_DG_0: Grupo no diabéticas gestacionales visita T0, Non_DG_T1: grupo no diabéticas gestacionales en la visita T1.

Validación de los DMPs en la cohorte EPIPREG

Los principales resultados se validaron en la cohorte independiente EPIPREG. Las características de estas participantes se muestran en la tabla 4.

Para la obtención de los DMPs en la cohorte EPIPREG se realizó un estadístico t moderado de eBayes con el paquete limma, ajustado para las mismas variables que en el análisis principal de la cohorte EPIDG: edad, IMC pregestacional, sexo fetal, edad gestacional y HOMAIR. En EPIPREG, el análisis se ajustó además por etnia: europea (EPIPREG_EU) y sudasiática (EPIPREG_SA). Se consideró estadísticamente significativo un valor de P <0,05.

En EPIPREG, intentamos replicar los 20 DMPs más significativos y los 3 DMPs que mejor discriminaba entre los grupos DG vs no-DG. Cg04095097

se replicó en la muestra completa ($p = 0.004$), y esto persistió después del ajuste por covariables (Tabla 7). De los 20 DMPs, 2 sitios CpGs presentaron diferencias estadísticamente significativas entre las mujeres con DG y sin DG en la muestra completa de EPIPREG (cg27603605 y cg12080079) (Tabla 7). Al igual que en la cohorte EPIDG, los niveles de metilación del ADN de estos CpGs fueron más altos en las mujeres embarazadas con DG en comparación con las mujeres sin DG (Figura 4). Al estratificar por etnicidad, observamos diferencias estadísticamente significativas solo para cg04095097 en las mujeres embarazadas de ascendencia surasiática (0.354 no DG vs 0.397 DG; $p = 0.026$), y una tendencia en las mujeres de ascendencia europea (Tabla suplementaria 4). Finalmente, un modelo de regresión logística mixto mostró que las mujeres embarazadas con niveles más altos de metilación de cg04095097 tenían una mayor probabilidad de desarrollar DG en comparación con las mujeres con niveles de metilación más bajos en la muestra de EPIPREG (OR = 1.25; IC del 95% = 1.01-1.52).

Tabla 7- CpGs validados en la cohorte EPIPREG ajustado por edad, IMC previo al embarazo, edad gestacional, HOMAIR, sexo fetal y ascendencia.

	logFC	Media non-DG	Media DG	P.Val
cg04095097	0.16912107	0.3358	0.3613	0.00424159
cg27603605	0.07076299	0.669	0.681	0.01647897
cg12080079	0.09568801	0.728	0.743	0.04782572
cg15329406	-0.0089716	0.29	0.288	0.80083925
cg01459453	-0.0147631	0.751795	0.7527223	0.81069312

LogFC: log₂ Fold Change, Mean Controls: media de los β -value en el grupo control. Mean DG: media de β -value en el grupo DG.

Por último, cuando realizamos un análisis de curva ROC con estos sitios CpGs en la cohorte EPIPREG, obtuvimos AUC estadísticamente

significativos, aunque con un AUC menor que en EPI-DG, para los siguientes sitios CpGs: cg04095097 (AUC = 0.570, $p = 0.015$), cg04095097 + cg27603605 (AUC = 0.597, $p = 0.0009$) y una combinación de los tres CpGs (AUC = 0.605, $p = 0.0003$) (Figura suplementaria 6).

mQTLs

A partir de búsquedas en GoDMC, encontramos 21 mQTLs para los 3 sitios CpG replicados. Según Phenoscanner, el mQTL para cg12080079, rs585075, estaba asociado con DG ($p = 0.03$). Del mismo modo, los otros mQTLs estaban nominalmente relacionados con DM2, variables metabólicas, y procesos inflamatorios autoinmunes como la hipertensión, la psoriasis, la enfermedad de Crohn o la artritis reumatoide (Tabla 8).

Tabla 8- mQTLs relacionados a los 3 CpGs replicados en EPIPREG

ID_cg	Gene	rs	Gene	Posición	A1	A2	Enfermedad (PMID)	Beta	P-val
cg12080079	PAX7	rs10907325	PAX7	chr1:18964021	A	T	Systolic blood pressure (PMID: 19609347)	NA	2.02E-05
							Body mass index within family tests max measurement (PMID: 21701565)	NA	2.50E-04
							Amyotrophic lateral sclerosis (PMID: 28931804)	-0.0112	7.33E-04
							Non-insulin dependent diabetes mellitus (PMID: UKBB)	0.0002683	6.98E-03
		rs115326690		chr1:19036153	C	G	Systolic blood pressure (PMID: UKBBq)	-0.02669	6.33E-04
							Hypertension (PMID: UKBB)	-0.01019	2.57E-03
							Psoriasis (PMID: UKBB)	-0.000736	2.76E-03
		rs12563020		chr1:19022599	A	G	Type II diabetes adjusted for BMI (PMID: 28566273)	0.099	4.60E-03
							Type II diabetes (PMID: 28566273)	0.064	2.70E-02
		rs585075		chr1:19021373	T	C	Gestational diabetes only (PMID: UKBB)	-0.01402	2.56E-02
rs7527312	chr1:19014865	C	G	Body mass index adjusted for smoking in males (PMID: 28443625)	0.0221	3.18E-03			
cg04095097		rs10906900	FAM171A1	chr10:15472860	A	G	High grade serous ovarian cancer (PMID: 28346442)	-0.05032	2.84E-03
		rs111256448	ITGA8	chr10:15765261	C	T	Crohns disease (PMID: 26192919)	-0.1337	1.27E-03
							Diastolic blood pressure (PMID: UKBB)	-0.01169	5.94E-03
		rs111920126	ITGA8	chr10:15510921	A	G	Clear cell ovarian cancer (PMID: 28346442)	-0.492	9.00E-04
		rs11259597	FAM171A1	chr10:15332911	A	G	Crohns disease (PMID: 26192919)	-0.211	3.73E-03
							Spondylosis (PMID: UKBB)	0.001257	4.26E-03
		rs11259690	ITGA8	chr10:15521768	A	G	Heart attack (PMID: UKBB)	-0.002578	1.36E-03
							Type 2 diabetes (PMID: UKBB)	-0.000909	3.31E-02
rs114410649	ITGA8	chr10:15555993	A	G	Non-insulin dependent diabetes mellitus (PMID: UKBB)	0.0006851	3.28E-02		

	rs117164512	FAM171A1	chr10:15298845	C	T	<i>Alcoholic hepatitis (PMID: UKBB)</i>	-0.004734	3.62E-03
	rs12266863	ITGA8	chr10:15526023	A	G	<i>Rheumatoid arthritis (PMID: 24390342)</i>	-0.08338	8.20E-03
	rs12355715	ITGA8	chr10:15534600	C	T	<i>Polyarthrosis (PMID: UKBB)</i>	-0.000452	3.96E-03
	rs137882390	FAM171A1	chr10:15265022	A	G	<i>Multiple valve disease (PMID: UKBB)</i>	0.0006208	1.12E-03
						<i>Illnesses of father: diabetes (PMID: UKBB)</i>	-0.008409	6.48E-03
	rs187688289	FAM171A1	chr10:15471793	C	T	<i>Low grade serous ovarian cancer (PMID: 28346442)</i>	-0.6751	7.10E-04
	rs61514659	FAM171A1	chr10:15467588	C	T	<i>Type 2 diabetes (PMID: UKBB)</i>	-0.002412	6.07E-04
	rs7087034	ITGA8	chr10:15738818	C	T	-		
	rs71493225	ITGA8	chr10:15536804	C	T	<i>Ischemic cardiomyopathy (PMID: UKBB)</i>	0.002818	3.17E-04
						<i>Insulin-dependent diabetes mellitus (PMID: UKBB)</i>	0.000251	2.98E-02
	rs77070234	RP11-455B2.9	chr10:15226599	A	C	<i>Eye problems or disorders: diabetes related eye disease (PMID: UKBB)</i>	0.003402	8.50E-03
	rs7915524	FAM171A1	chr10:15293903	C	T	<i>Self-reported mitral regurgitation or incompetence (PMID: UKBB)</i>	0.0001378	2.50E-03
						<i>Insulin-dependent diabetes mellitus (PMID: UKBB)</i>	-0.000197	4.12E-02
cg27603605		Non-SNPs						

UKBB: United Kingdom BioBank (www.ukbiobank.ac.uk). A1: allele dominant, A2: allele recessive.

Objetivo 2- Comparación del perfil epigenético en muestras de saliva entre la descendencia nacida de madres con DG y la de madres sin DG.

Epigenome wide association study identifies CpGs associated with the growth pattern in offspring from mothers with gestational diabetes mellitus: EPIDG cohort.

The aim of this objective is to identify epigenetic marks in infant saliva that differentiate between children of mothers with DG and children of non-DG mothers. Additionally, to determine whether these marks persist in the first year of life and whether are associated with the growth pattern of these children in the postpartum period and one year after birth. An epigenome wide association study (EWAS) from saliva samples was made in a total of 28 children from EPIDG cohort, 15 of whom were born from non-DG women, and the other 13 from women with DG. Saliva samples were taken two weeks after birth (T.2) and one year later (T.3). Differentially methylated positions (DMP) were obtained using limma package for R studio, adjusted by newborn sex, DG treatment (diet or insulin), and maternal weight gain ($P.val < 0.01$). A total of 5.132 DMPs were found in the postpartum period, and 6.960 DMPs at one year after birth. 16 of these DMPs were common in both points (T2 and T3) in the same direction. A correlation analysis with these DMPs, showed that methylation values in T2 of 4 CpGs were corelated with weight, head circumference and height percentiles. The methylation values at T3 of 6 CpGs were also related with those percentiles. In conclusion, we have found a differentiated DNA methylation pattern in saliva samples between children born from mothers who suffered DG during pregnancy compared to children born from non-DG mothers. Some of these DMPs persist at one year of age and are associated with variables related to the child's growth pattern.

Características de la cohorte

	Tipo de muestra	1ª visita post parto (8 semanas)	1 año pos parto
Descendencia de madres DG	Saliva	15	15
Descendencia de madres no DG		15	15

Para este array se seleccionaron a aquellas participantes que acudieron con sus hijos a las visitas posparto y al año, de ambos grupos (DG y no DG). Se obtuvieron 15 muestras de controles y 15 de nacidos de madres con DG. Tras pasar los controles de calidad del ADN, se redujeron a un total de 56 muestras.

Las características de estos sujetos se muestran en la Tabla 9. No se encontraban diferencias significativas en las variables antropométricas de los niños en el posparto comparando los niños nacidos de madres con DG frente a los nacidos de madres sin DG. Al año del nacimiento se vio que había diferencia en el peso, pero a los dos años esta diferencia dejaba de ser significativa, siendo el peso igual en los dos grupos.

Tabla 9- Característica de las variables antropométricas de los niños incluidos en el array en el momento de posparto, un año de vida y a los dos años de vida.

Visita	Recién Nacido			1 año			2 años			
	Grupo	no DG	DG	P.val	no DG	DG	P.val	no DG	DG	P.val
Percentil de Peso		59.27±7.28	60.83±8.55	0.864	67±6.68	58.31±9.04	0.596	57.91±9.55	63.83±9.43	0.622
Percentil de Talla		50.87±6.56	59.58±8.55	0.393	50.61±7.2	53±10.15	0.885	50.83±14.37	59.13±14.18	0.518
Percentil Perímetro cefálico		43.47±6.44	38.67±9.66	0.406	63.53±6.11	55.25±10.63	0.751	57.73 ±8.77	65.4±7.96	0.457
Sexo (niño/niña)		6/9	6/9	0.653						
Tipo de parto (vaginal/cesárea)		9/6	9/6	0.558						

Los datos están expresados como la media ± desviación estándar. Las proporciones fueron comparadas mediante test chi-cuadrado y las variables cuantitativas fueron analizadas usando t-test o U de Mann-Whitney dependiendo de la normalidad de los datos.

Análisis de metilación diferencial

Se encontraron un total de 6,968 DMPs y 5,132 DMPs en el posparto y después de un año de seguimiento, respectivamente. Los CpGs más diferenciados por cromosoma en la visita posparto se representan en el Manhattan plot de la figura 16 y al año en la figura 17. La inflación en el caso del posparto λ fue 1, lo que significa que no hubo inflación, y encontramos una inflación moderada en el análisis de seguimiento de un año ($\lambda = 1.36$) (figura suplementaria 7).

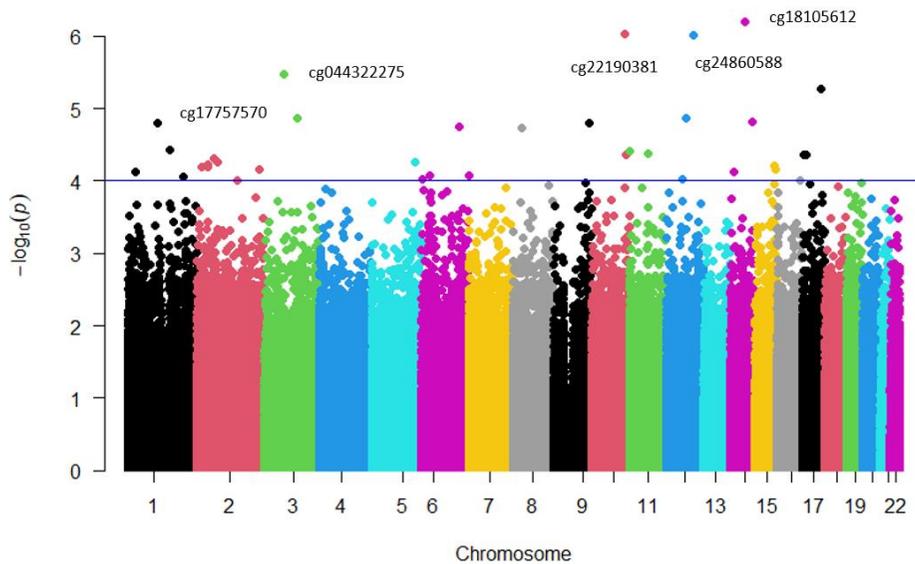


Figura 16-Manhattan plot que representa los CpGs más diferenciados por cromosoma en la visita inicial.

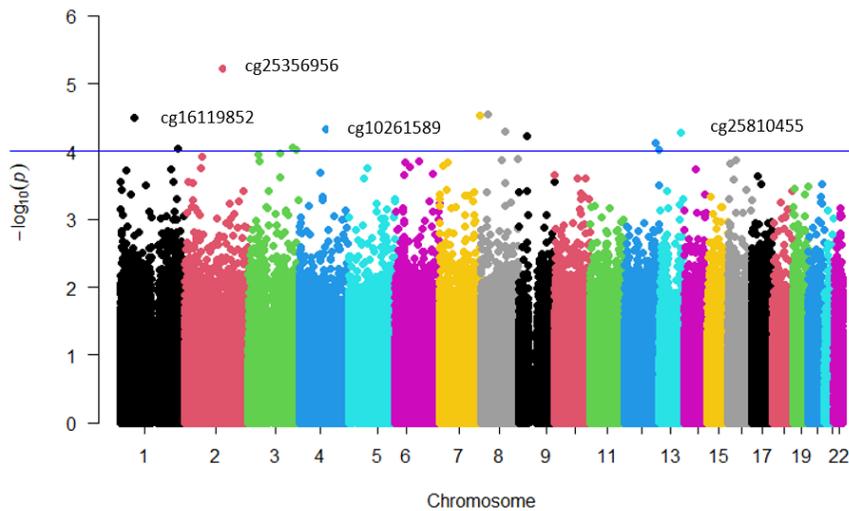


Figura 17- Manhattan plot que representa los CpGs más diferenciados por cromosoma en la visita preparto.

Un total de 50 DMPs se mantuvieron diferenciados entre ambos grupos a lo largo del tiempo. Sin embargo, solo 16 CpGs mantuvieron la misma dirección (Figura 18).

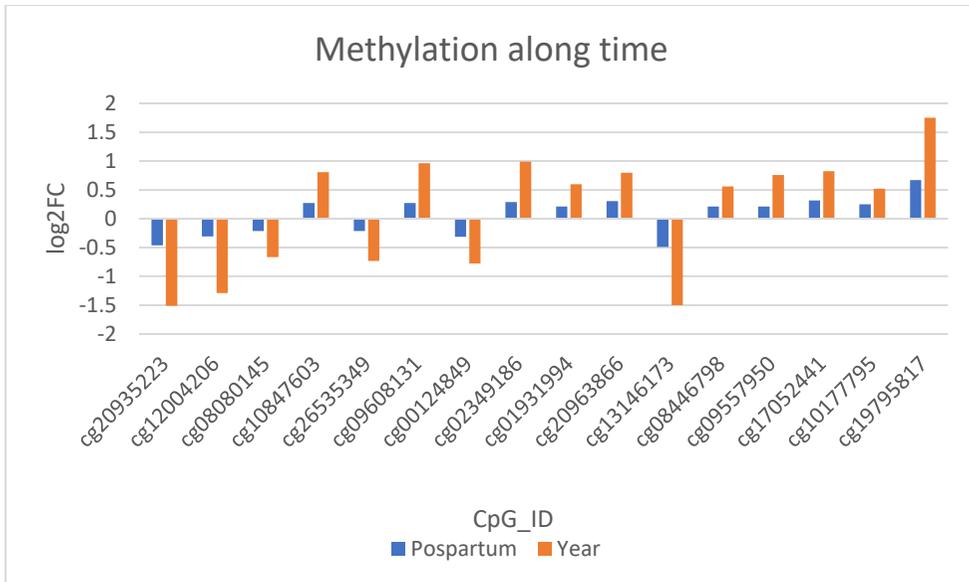


Figura 18- 16 CpGs que mantienen la misma dirección en cuanto a la metilación del ADN a lo largo del año. Pospartum: visita posparto, Year: visita al año.

Análisis enriquecimiento funcional

De los 16 DMPs diferenciados en ambos períodos, solo 8 se encontraban anotados en genes. Utilizamos esas 8 proteínas para realizar un análisis funcional con String y ENRICH. La red resultante de String con las opciones predeterminadas no mostró ninguna conexión entre las proteínas. Con el fin de obtener una red con las proteínas que interactuarían con nuestras proteínas de interés, se configuró STRING para proporcionar una red con un número máximo de interactores de 100. La red resultante (figura suplementaria 8) se cargó en Cytoscape para realizar una clusterización con la aplicación MCODE. Cytoscape devolvió 4 grupos, solo dos de ellos incluyeron algunas de nuestras proteínas de interés.

El primer grupo incluyó DEFB104A, y estuvo relacionado con la infección por *Staphylococcus aureus* y la vía de señalización del receptor NOD-like, que está involucrada en la respuesta a la infección. FARP2 y CYTH3 estaban

en el segundo grupo. El estudio de KEGG en este grupo se vio que estaba relacionado con diferentes vías, siendo las más interesantes la vía de señalización de fosfolipasa D, la vía de señalización del VEGF (relacionada con el proceso inflamatorio) y, aunque menos significativa, también encontramos que el grupo estaba relacionado con la diabetes tipo 2 (figura 19).

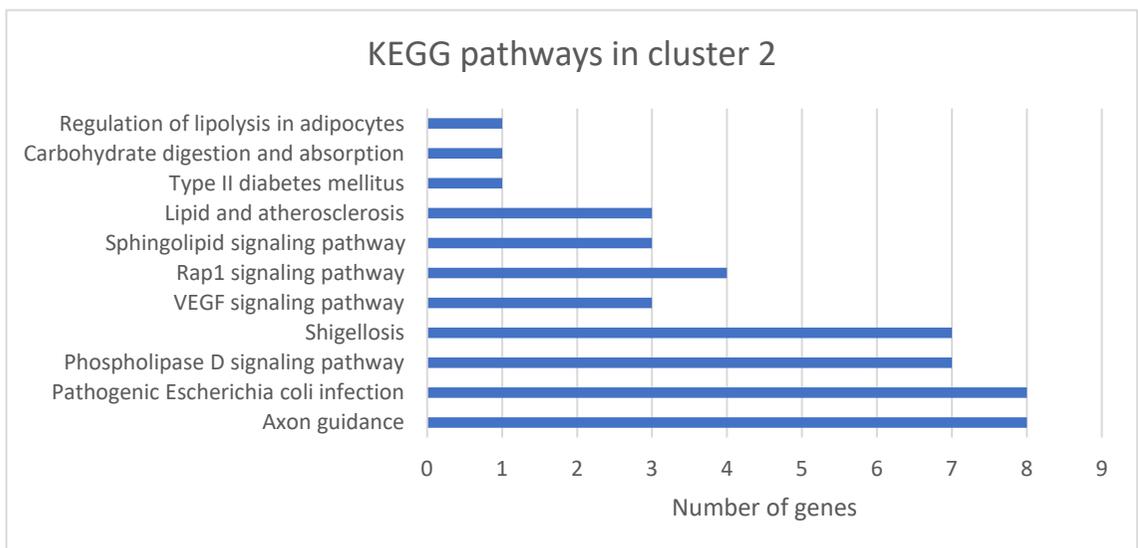


Figura 19- Rutas de señalización más enriquecidas según KEGG.

Análisis de correlación y regresión

Para el análisis con variables antropométricas de los niños, se utilizaron los percentiles según la edad y el sexo, siguiendo las recomendaciones de la OMS. Se realizaron correlaciones con los percentiles de peso, talla y perímetro de la cabeza en el momento del nacimiento, en la visita al año y a los dos años. Observamos que el percentil del perímetro cefálico al año se correlacionaba con los siguientes CpGs: cg00124849 y cg12004206 en el posparto, cg10177795, cg01931994 y cg10847603 al año. El percentil de peso del recién nacido y al año correlacionaba con el nivel de metilación

del cg01931994. También vimos que el percentil del peso al año se correlacionaba con el cg02349186 en el posparto, y el percentil de peso a los dos años se correlacionaba con 3 CpGs al año (cg10177795, cg19795817 y cg20963866) (Figura 20B). Por último, se obtuvo que el percentil de la talla a los dos años se asoció con el cg19795817 tanto posparto como al año, y con los CpGs siguientes: cg00124849, cg10177795, cg20963866 (Figura 20C). El cg01931994 se encontraba anotado en el gen FAM5B y el cg10177795 en el gen FARP2. El resto de CpGs correlacionados no tenían genes anotados.

Búsqueda de CpGs relacionados con la obesidad infantil

Se realizaron búsquedas en la base de datos de Phenoscanner (<http://www.phenoscanner.medschl.cam.ac.uk/>) para identificar genes previamente asociados por GWAS a la obesidad infantil, utilizando un umbral de valor de p $1E-5$. Se encontraron un total de 32 genes, en los cuales se encontraban un total de 877 CpGs en la base de datos de anotaciones de Illumina (tabla suplementaria 5).

Se extrajeron los M y β -values para ver si existían diferencias de metilación entre los grupos. Se encontraron 41 DMPs en el posparto y 32 DMPs en el seguimiento de un año. Dos De estos DMPs fueron comunes en ambos momentos (cg27107076, cg03641192).

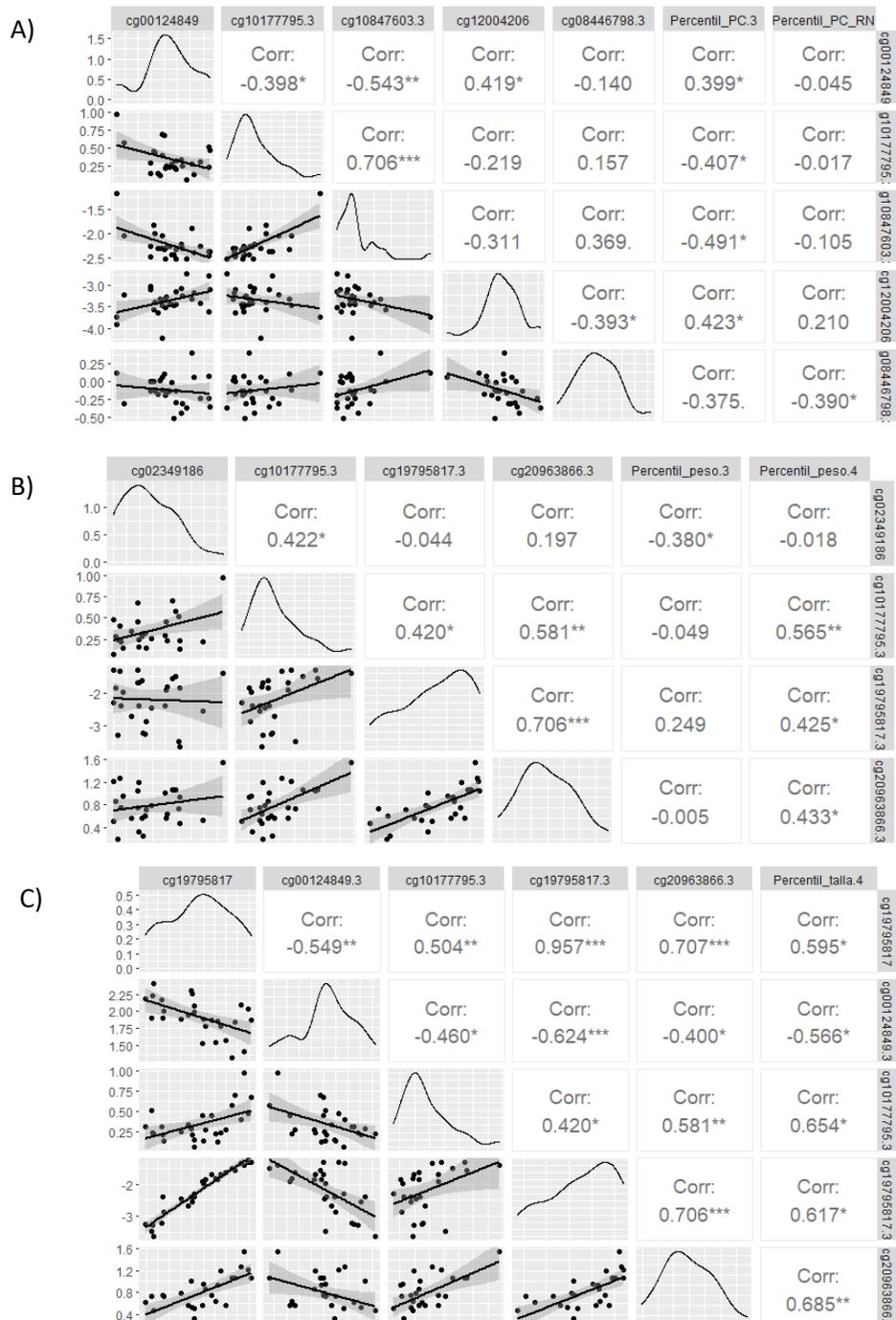


Figura 20- Correlaciones de los valores de los CpGs con los percentiles de los niños. A) Correlación de los CpGs significativos con el percentil del perímetro cefálico en distintos tiempos. B) Correlación de los CpGs significativos con el percentil del peso en distintos tiempos C) Correlación de los CpGs significativos con el percentil de la talla en distintos tiempos. PC: Percentil cefálico, RN: recién nacido, .3: visita al año, .4: visita a los dos años. *: $pval < 0.05$, **: $pval < 0.005$, ***: $pval < 0.0005$.

Objetivo 3- Evaluación de la relación de un MRS de DM2 con la DG

DNA methylation risk score for type 2 diabetes is associated with gestational diabetes.

In this objective, we aimed to test whether a DNA methylation risk score (MRS) for T2DM was associated with DG across ancestry and DG criteria. In two independent pregnancy cohorts, EPIPREG (n=480) and EPIDG (n=32), DNA methylation in peripheral blood leukocytes was measured at a gestational age of 28+-2. We constructed an MRS in EPIPREG and EPIDG based on CpG hits from a published epigenome-wide association study (EWAS) of T2DM. With mixed models logistic regression of EPIPREG and EPIDG, MRS for T2DM was associated with DG: odd ratio (OR)[95% CI]: 1.3 [1.1-1.8], P=0.002 for the unadjusted model, and 1.4 [1.1-1.7], P=0.00014 for the model adjusted by age, pre-pregnant BMI, family history of diabetes and smoking status). Also, we found 6 CpGs (cg14020176, cg22650271, cg14870271, cg27243685, cg06378491, cg25130381) associated with DG, and some of their methylation quantitative loci (mQTLs) were related to T2DM and DG. In conclusion, for the first time, we show that DNA methylation marks for T2DM are also associated with DG, suggesting shared epigenetic mechanisms between DG and T2DM.

Característica de las muestras

Las características clínicas de las cohortes EPIPREG y EPIDG se representan en la Tabla 4. En la cohorte EPIPREG, las mujeres no-DG y DG difirieron significativamente en el IMC pregestacional, la glucosa en ayunas (Gluc0), la glucosa a las 2 horas después de la prueba de tolerancia oral a la glucosa (Gluc2), la presión arterial sistólica (PAS), la hemoglobina A1C glucosilada (HbA1C), el Modelo de Evaluación de la Homeostasis de Insulina (HOMAIR), los triglicéridos (TG), el colesterol y el HDL en todas las etnias. En la figura 21 se detalla el flujo de trabajo establecido para la selección de los grupos.

EPIPREG	Tipo de muestra	Europeas	Sudasiáticas
Embarazada DG	Sangre	73	68
Embarazada no DG		239	100

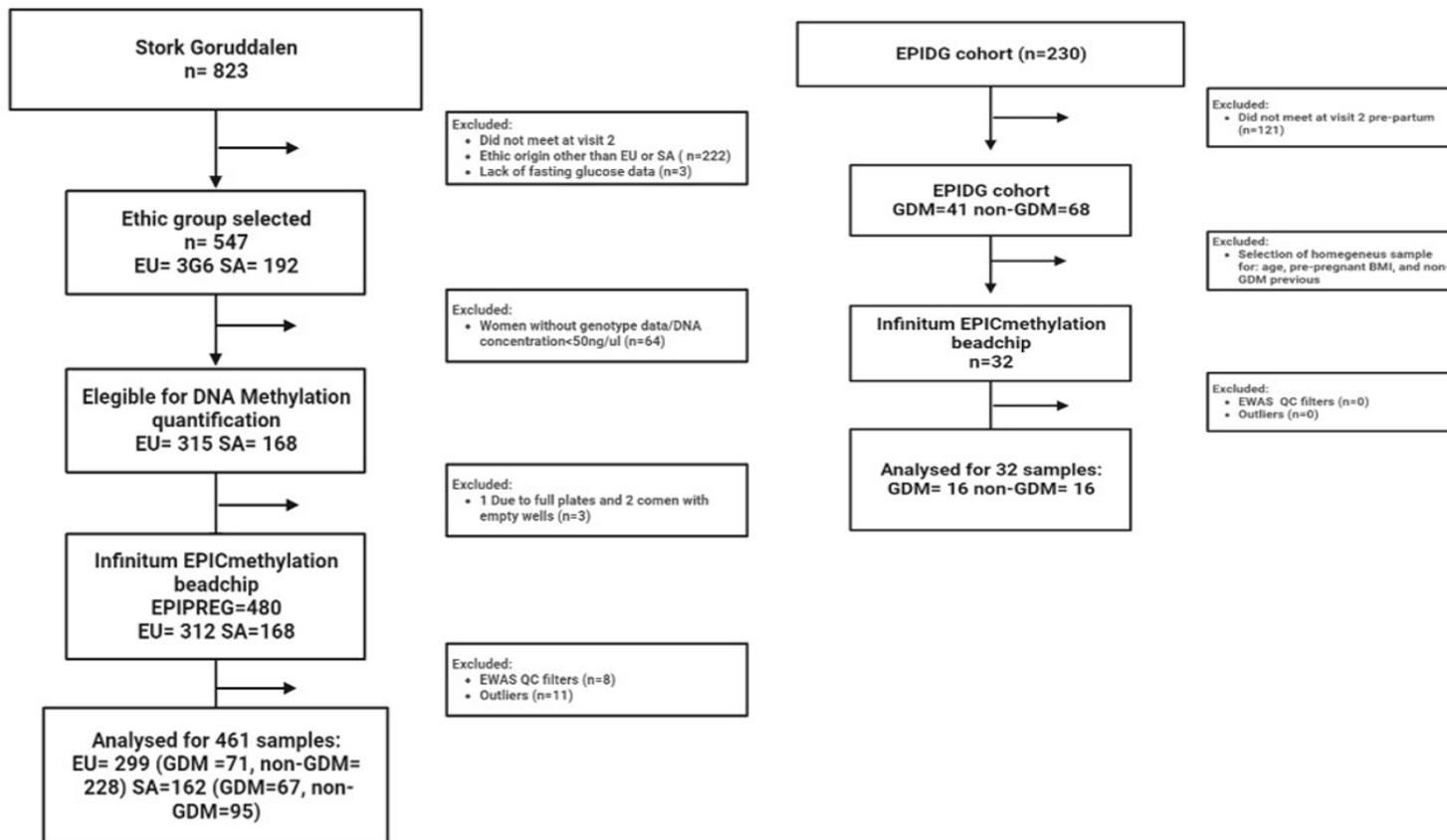


Figura 21- Flujo de trabajo de selección de las muestras utilizadas para realizar la asociación con el MRS de diabetes tipo 2

Cálculo del MRS

Una vez calculados los MRS para cada cohorte, observamos como el MRS es mayor en el grupo de DG en las 3 poblaciones (figura 22). Como se menciona en la introducción, la construcción de un MRS hay que identificar los CpGs asociados a un fenotipo de interés y usar el peso del efecto sobre dicho fenotipo. El valor del peso del CpG se multiplica por el β -value de ese CpG en nuestra cohorte en cada paciente y se realiza el sumatorio de ese CpG para cada paciente (ver figura 7- introducción). Además, se puede apreciar como los niveles de metilación son muy diferentes entre cohortes.

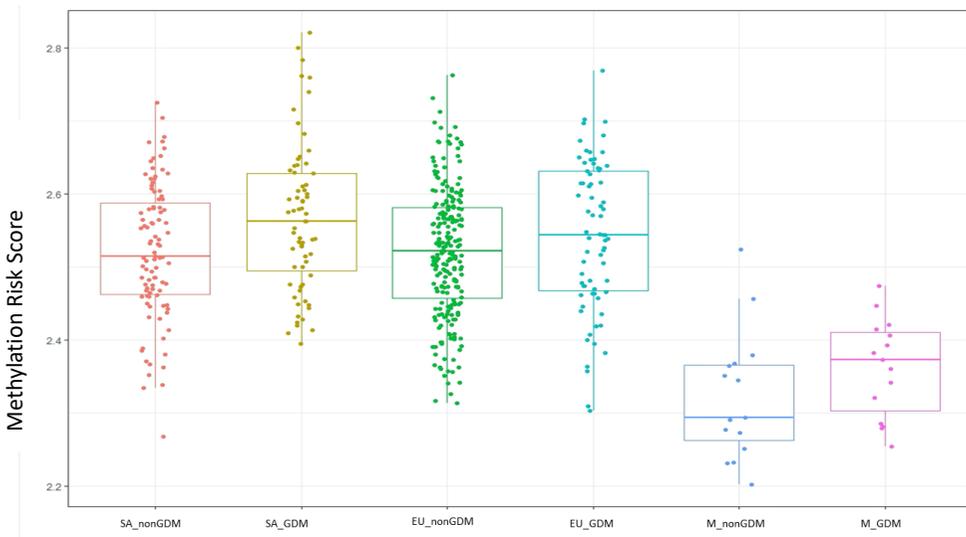


Figura 22- Boxplot representa el MRS en cada paciente separado por grupo al que pertenece. Rojo (SA_nonDG): grupo EPIPREG_SA no diabéticas, Amarillo (SA_DG): EPIPREG_SA grupo diabéticas gestacionales, Verde (EU_nonDG): EPIPREG_EU no diabéticas, Azul claro (EU_DG): EPIPREG_EU grupo de diabéticas, Azul oscuro (M_nonDG): EPIDG grupo no diabéticas, Rosa (M_DG): EPIDG grupo diabéticas.

Para analizar la asociación entre el MRS y la DG realizamos modelos mixtos de regresión logística, observando que un mayor MRS se asociaba con una mayor probabilidad de tener DG, incluso cuando ajustamos por la edad, IMC gestacional, antecedentes familiares de diabetes y el tabaco (O.R: 1.4, 95%C.I: 1.10-1.74, P.val= 0.00014). En cuanto a las variables continuas,

tanto la glucosa en ayunas como el péptido C se asociaron de forma significativa con el MRS mediante un modelo mixto lineal. Sin embargo, cuando ajustamos por las mismas variables que en el modelo anterior solo la glucosa en ayunas mantuvo su asociación de forma estadísticamente significativa (P.val=0.0086) (tabla10).

Tabla 10- Modelos mixto de regresión lineal del MRS con las variables continuas

	Modelo	Estimate	Std.Error	df	t.value	P.val
HOMAIR		0.989	0.802	479	1.23	0.218
C-peptide	No ajustado	0.422	0.167	0.0483	2.535	0.0116
Gluc0		0.8351	0.2726	0.0279	3.064	0.0024
Pre-IMC		4.36	2.11	0.0316	2.06	0.04
HOMAIR		Ajustado por edad	0.416	0.777	0.0468	0.536
C-peptide	IMC previo, antecedentes y consumo tabaco	0.254	0.152	4.780	1.67	0.096
Gluc0		0.707	0.268	0.0313	2.641	0.00869
Pre-IMC		Ajustado por edad, antecedentes y consumo tabaco	3.72	2.04	1.15	1.83

C-peptide: Péptido C, Gluc0: Glucosa en ayunas, Pre-IMC: IMC Previo al embarazo

mQTLs para los sitios CpG comunes entre DM2 y DG

Para identificar los CpGs que se encuentran asociados a la GDM de forma individual en cada grupo, se realizó una regresión logística para cada CpG usado en el MRS en cada uno de los 3 grupos. Tras ello, se realizó un metaanálisis para ver que CpGs se relacionaba en los tres grupos con la DG. En el metaanálisis de los 42 sitios CpG incluidos en el MRS para DM2, seis CpGs estuvieron significativamente asociados con DG (FDR<0.05) (tabla 11).

Tabla 12-1- Lista de 6 CpGs significativamente asociado con DG en las tres cohortes (EPIPREG_EU, EPIPREG_SA, EPIDG).

CpG	Effect	StdErr	Dirección	FDR	Genes
cg14020176	1364.7387	334.8719	+++	0.00192906	SLC9A3R1
cg22650271	2288.3955	612.9638	+++	0.003969	SYNGR1
cg14870271	10.0163	2.7782	+++	0.0043652	LGALS3BP
cg27243685	2434.2168	716.9641	+++	0.00576156	ABCG1
cg06378491	3200.9735	938.6366	+++	0.00576156	MAP4K2
cg25130381	106813769	38646.6933	+++	0.039984	SLC9A1

Todos los sitios CpG identificados excepto cg25130381, tuvieron mQTLs significativos (tabla 12). Se encontraron un total de 23 mQTLs. Según las búsquedas en Phenoscanner, estas variantes genéticas estaban nominalmente asociadas con fenotipos relevantes como DG, T2DM, enfermedad coronaria y enfermedades inflamatorias, pero no se han descubierto en estudios de asociación a nivel genómico debido a que no alcanzaron la significancia a nivel genómico (p que oscilan entre 0.000002 y 0.05). En búsquedas de datos resumidos de GWAS utilizando Phenoscanner, el rs74374453 en AP005273.1, mQTL para cg06378491 en MAP4K2, estuvo asociado con DG (P.val= 0.00299), con el factor de crecimiento similar a la insulina 1 (P.val=0.049) y con la diabetes mellitus tipo 1 (T1DM) (P.val =0.043). El rs7289325 en el gen SYNGR1, mQTL de cg22650271, estuvo asociado con DG (P.val= 0.042), y con el tratamiento con insulina (P.val=0.0125). Además, varios de nuestros mQTLs identificados estuvieron asociados con DM2 en datos resumidos de GWAS: rs2385067 (P.Val= 0.00046), rs652963 (P.val= 0.028), rs5757582 (P.val= 0.0094), rs742402 (P.val= 0.0095) y rs9611137 (P.val= 0.029) (tabla 12).

Tabla 12- mQTLs and Disease associated, according to Phenoscaner.

ID_cg	Gene of cg	Chr	Position	rs	Gene of rs	Chr	Posicion		A1	A2	Disease	Beta	P-val
cg06378491	MAP4K2	11	64564012	rs74374453	AP005273.1	11	64271540	cis	A	G	Insulin-like growth factor 1		4.96E-02
											Gestational diabetes only	-0.086	2.99E-03
											Self-reported type 1 diabetes	-0.0007011	4.28E-02
cg14020176	SLC9A3R1	17	72764985	rs2385067	TMEM104	17	72810070	cis	A	G	Type II diabetes	-0.195	4.62E-04
				rs652963	QSER1	11	32911737	trans	T	C	Illnesses of siblings: diabetes	-0.002764	7.23E-03
											Type II diabetes	-0.036	2.80E-02
cg14870271	LGALS3BP	17	76976010	rs117549034	USP36	17	76826416	cis	C	T	Parkinson Disease	-0.02651	2.82E-06
											Diabetes mellitus in pregnancy	-0.0004308	1.18E-02
											Started insulin within one year diagnosis of diabetes	0.04261	1.40E-02
											Eye problems or disorders: diabetes related eye disease	-0.005733	4.11E-02
cg22650271	SYNGR1	22	39760165	rs5757582	AL031590.1	22	39661032	cis	A	C	Rheumatoid arthritis	-0.06188	1.90E-05
											Crohns disease	0.000472	4.76E-04
											Type II diabetes adjusted for BMI	0.038	9.40E-03
				rs7289325	PDGFB	39642577	cis	A	T	Started insulin within one year diagnosis of diabetes	-0.009563	1.25E-02	
										Gestational diabetes only	0.01211	4.44E-02	
				rs742402	AL031590.1	39659487	cis	A	G	High cholesterol	-0.002935	4.90E-04	
										Crohns disease	-0.0004341	2.30E-03	
Rheumatoid arthritis	0.06766	1.70E-05											

											Type II diabetes adjusted for BMI	-0.04	9.50E-03
											Illnesses of mother: diabetes	-0.001557	4.31E-02
											Medication for cholesterol, blood pressure or diabetes: none of the above	0.003574	4.64E-02
			rs9611137	SCUBE1		39682445	cis	C	T		Crohns disease	-0.2224	3.76E-05
											Inflammatory bowel disease	-0.152	8.48E-05
											High cholesterol	-0.00366	2.90E-02
											Medication for cholesterol, blood pressure or diabetes: cholesterol lowering medication	-0.00794	1.33E-02
											Type II diabetes adjusted for BMI	-0.071	2.90E-02

Discusión

El patrón de metilación del ADN en sangre periférica de mujeres gestantes con y sin DG es diferente y afecta a numerosos genes asociados a la DMT2

La mayoría de los estudios epigenéticos sobre la diabetes gestacional se han realizado en muestras de placenta y sangre del cordón umbilical (123). Solo unos pocos estudios han evaluado el patrón epigenético en muestras de sangre de mujeres embarazadas. Wu et al. evaluaron cambios en el epigenoma de mujeres embarazadas antes del desarrollo de DG mediante el array 450K de Illumina, en muestras de sangre materna (124). Identificaron 5 CpGs (asociados a los genes *COPS8*, *PIK3R5*, *HAAO*, *C5orf34* y *CCDC124*) como posibles biomarcadores clínicos para la detección temprana de DG y la intervención terapéutica. Estos autores validaron los resultados mediante pirosecuenciación, encontrando una buena correlación con los valores β de metilación, pero no replicaron los hallazgos en otras cohortes. Dias et al. examinaron la relación entre la DG y el perfil de metilación del ADN en las células sanguíneas periféricas de mujeres sudafricanas (125). Compararon el epigenoma mediante el array de Illumina EPIC en mujeres embarazadas con DG y sin DG en el primer trimestre. Al igual que el estudio de Wu et al., varios sitios CpGs mostraron metilación diferencial, y los genes asociados estaban involucrados en vías de regulación metabólica. Además, los cinco sitios CpG principales se correlacionaron con variables metabólicas relevantes, como los niveles de glucosa e insulina. Kang et al. analizaron el epigenoma en muestras de sangre materna y de cordón umbilical de 16 mujeres embarazadas (8 con DG) y sus descendientes (126). Encontraron un patrón de metilación diferencial entre el grupo con DG y el grupo no expuesto en ambos tipos

de muestras, respaldando la hipótesis de que la DG tiene un efecto epigenético tanto en las madres como en sus descendientes.

Los resultados obtenidos en este trabajo corroboran estudios previos en los que se observa un patrón de metilación diferencial entre DG y controles. Nuestros resultados son más amplios que trabajos anteriores y, aunque coinciden en muchos aspectos, han permitido precisar mejor las vías metabólicas y los mecanismos implicados.

Del total de 1041 DMPs encontrados en nuestro estudio, 714 están en genes anotados (la anotación consiste en la identificación de todas las características relevantes de una secuencia del genoma), de los cuales encontramos 23 genes relacionados con la DMT2. Algunos de los genes se correlacionaron con variables bioquímicas relacionadas con la DG (glucosa en ayunas y en diferentes puntos de la SOG, HbA1C, HOMAIR). Los patrones de metilación del ADN encontrados afecta a muchos otros procesos, entre los que cabe citar la transducción de señales, el crecimiento celular, la proliferación, la diferenciación y apoptosis celular, resistencia a la insulina, metabolismo de la glucosa, inflamación, señalización neurológica y oncogénesis (124,127,128). De acuerdo con nuestro análisis de enriquecimiento funcional de los DMPs, encontramos que las GO y KEGG más significativas estaban relacionadas con la regulación de la secreción de insulina y la vía de señalización de la misma, dos aspectos clave en la DMT2 (129–131).

CAMTA1 fue uno de los genes más hipometilados que hemos encontrado en pacientes con DG. *CAMTA1* parece estar implicado en la regulación y secreción de insulina de las células beta (132). La hipometilación observada

podría implicar una expresión alterada de *CAMTA1* que afecta la secreción de insulina y, en consecuencia, la regulación del metabolismo de la glucosa y los lípidos (133), tal y como observamos con una correlación negativa entre algunos CpGs de este gen y el HOMA-IR y los niveles de triacilglicéridos. Dias et al. encontraron, como nosotros, una menor metilación de *CAMTA1* en mujeres con DG (128). En ambos casos la ubicación en el gen de los sitios CpGs estudiados fue similar. Finalmente, diferentes estudios han mostrado una asociación entre polimorfismos en el gen *CAMTA1* y el metabolismo de los carbohidratos, lípidos en sangre, presión arterial (134), glucosa en ayunas (135) y DMT2 (136).

Otros de los CpGs que encontramos, están en genes que parecen estar asociados con la respuesta a la sobrecarga oral de glucosa. Algunos de los DMPs encontrados se observan en la quinasa AKT serina/treonina 2 (*AKT2*) y la quinasa diana mecánica de rapamicina (*MTOR*). Distintos trabajos apuntan a que la regulación negativa de *AKT2* está relacionada con niveles elevados de glucosa y mayor resistencia a la insulina (134). En humanos, ciertos polimorfismos asociados con niveles disminuidos de *AKT2* se relacionaron con alteraciones en el metabolismo de la glucosa, resistencia a la insulina y DMT2 (135,136). Nosotros hemos encontrado que *AKT2* está hipermetilado en la región del promotor (5'UTR), lo que sugiere que este gen podría estar reprimido (137) y, por lo tanto, contribuir al aumento de los niveles de glucosa. En nuestro estudio, los CpGs de *MTOR* estaban hipermetilados en la región del cuerpo, lo que podría implicar una mayor expresión de este gen y esta sobreexpresión está asociada con procesos anabólicos y secreción de insulina (138,139). Finalmente, el CpG anotado en el gen *PRKAA2* se ha asociado en nuestro estudio con niveles elevados

de glucosa en ayunas y este gen parece estar relacionado con la sensibilidad a la insulina (140) y la DMT2 (141). Además, *PRKAA2* está estrechamente relacionado con la regulación de *MTOR*, de hecho, la regulación negativa de *PRKAA2* induce la expresión de *MTOR* y un aumento en la síntesis de proteínas y resistencia a la insulina en ratas (142). Quizás, una mayor metilación en la región TSS1500 de *PRKAA2*, como hemos encontrado, podría significar una regulación negativa de este gen, lo que resultaría en una disminución de la secreción de insulina y niveles de glucosa más altos después de la SOG.

En nuestros resultados, los genes *NBL1* y *MINOS1-NBL1* tenían más de 1 CpG correlacionado con la respuesta a la SOG a los 120 minutos, después del parto. *NBL1* parece estar implicado en la regulación de la función de las células beta y en la enfermedad renal diabética al actuar como antagonista de la proteína morfogenética ósea (BMP), que forma parte de la vía de señalización TGF- β (137,138). Por lo tanto, un nivel más bajo de *NBL1* podría inducir una disfunción en la función de las células beta. Además, hay estudios que sugieren la implicación de la familia BMP en la obesidad y el metabolismo de la glucosa (139). Por ello, el papel antagonista de *NBL1* a través de BMP podría estar implicado en el nivel de glucosa después de una SOG en la visita posparto.

Con el diseño de nuestro estudio no podemos demostrar que las asociaciones encontradas impliquen causalidad. No obstante, nuestros resultados coinciden con otros trabajos de la literatura que muestran un perfil epigenético diferente entre mujeres con DG y sin DG, y la implicación de algunos de estos genes en rutas metabólicas relacionadas con la DG o la DM2, u otras alteraciones metabólicas asociadas.

Las diferencias encontradas son relevantes y significativas más allá del umbral de FDR ($<0,05$). La probabilidad de falsos positivos es baja debido al pequeño factor de inflación genómica λ obtenido. En cuanto a las diferencias en los grupos, el grupo de referencia, no-DG, son mujeres embarazadas con una prueba de O'Sullivan positiva previa, por lo que, aunque se ha utilizado como control no debe extrapolarse a otras poblaciones. La principal diferencia entre los grupos fue que el grupo control mostró un mayor aumento de peso durante el embarazo en comparación con DG. Este hallazgo no es sorprendente e indica la buena adherencia de las mujeres embarazadas con DG a las medidas terapéuticas implementadas tras el diagnóstico.

Existe un patrón epigenético común de metilación a lo largo del embarazo que diferencia entre DG y no-DG.

El enfoque prospectivo del array nos ha permitido analizar por primera vez el perfil epigenético en muestras de sangre periférica de mujeres embarazadas con y sin DG en dos momentos durante el embarazo. El análisis de enriquecimiento funcional con aquellos DMPs que se mantenían diferencialmente metilado a lo largo del embarazo nos ha mostrado que algunos de ellos se relacionan con rutas del metabolismo de la glucosa.

Encontramos 272 DMPs en DG frente a no-DG en dos momentos durante el último trimestre del embarazo. Estos DMPs se encuentran en genes involucrados en la fisiopatología de la DG, como *SELP*, *LGR6*, *NBL1*, *RPS6KA2* o *ADCYAP1R1*. También detectamos tres DMPs, cg01459453, cg15329406 y cg04095097, que discriminaron con gran especificidad y

sensibilidad la ocurrencia de DG y replicamos uno de estos en una cohorte independiente.

El DMP cg01459453 se anotó al gen *SELP*. El gen *SELP* codifica la P-Selectina, una proteína de membrana granular y una molécula de adhesión celular que media la interacción de las células endoteliales o plaquetas activadas con los leucocitos. Varios estudios epidemiológicos sugieren que la disfunción endotelial está estrechamente relacionada con la resistencia a la insulina, precediendo al desarrollo de la DMT2 (140,141). El deterioro de esta función puede disminuir el paso de insulina al intersticio, limitando la acción de la insulina (141). Asimismo, otros autores han encontrado que los niveles elevados de P-selectina están asociados con el síndrome metabólico, y su expresión y secreción se han relacionado con bajos niveles de HDL-colesterol y una elevada glucosa en ayunas (142).

El DMP cg04802986, se anotó al gen del receptor acoplado a proteínas G con repeticiones ricas en leucina 6 (*LGR6*). En estudio anteriores, los sitios CpGs de este gen en el epigenoma de la descendencia, se asociaron con la presencia de DG en las mujeres embarazadas, así como con los niveles de glucosa en ayunas, a las 1 y 2 horas después de una prueba de tolerancia oral a la glucosa (SOG). Estas marcas epigenéticas se atenuaron mediante una intervención con dieta y actividad física durante el embarazo (143), lo que sugiere que el efecto de los niveles elevados de glucosa materna en la metilación del ADN en la descendencia podría modificarse mediante una intervención en el estilo de vida.

Igual que en el análisis de la metilación solo en la visita basal, el gen *NBL1* mostró que la metilación de los CpGs en las dos visitas diferenciaba entre

los dos grupos, por lo que estos cambios se mantienen a lo largo del embarazo. Tres de ellos se encontraban entre los 20 DMPs más significativos. Al igual que *SELP*, varios sitios CpG se han asociado con procesos relacionados con la edad (144) y la edad gestacional (145). Además, los sitios CpG de *NBL1* también se han asociado con el IMC materno previo al embarazo en un metaanálisis realizado en la cohorte de Epigenética del Embarazo y la Infancia (PACE, por sus siglas en inglés)(146).

El cg06279296 se encuentra en el gen *DIP2C*. Según la literatura, el gen *DIP2C* es muy susceptible de ser modificado por factores externos (147,148), por lo que la hiperglucemia durante el embarazo podría ser otro factor modificador.

Finalmente, nuestro análisis de enriquecimiento identificó dentro de nuestros DMPs, que se mantienen diferencialmente metilados a lo largo del embarazo, vías relevantes relacionadas con la Diabetes Mellitus Tipo I (*PTPRN2*), la resistencia a la insulina (*RPS6KA2*) y la secreción de insulina (*ADCYAP1R1*). Estos resultados han sido confirmados por varios estudios. Por ejemplo, *RPS6KA2* se ha relacionado con las vías de señalización de la insulina y las complicaciones de la DMT2 (149,150).

Potenciales biomarcadores epigenéticos de DG

Establecer biomarcadores capaces de identificar la ocurrencia de DG en el primer trimestre del embarazo es de enorme interés. Actualmente, el diagnóstico de DG se realiza con una prueba de tolerancia oral a la glucosa (SOG) alrededor de las 26-28 semanas de gestación. Encontrar biomarcadores accesibles y estables para predecir la DG permitiría una intervención temprana y la reducción subsiguiente del riesgo tanto en las madres como en sus descendientes.

Varios estudios han postulado algunos CpG como marcadores de DG. 62 sitios CpG específicos relacionados con DG, 6 de ellos ubicados en zonas promotoras, se incluyeron en un modelo predictivo para la DG, que obtuvo buenos resultados (151). Por su parte, Enchun et al., han identificado posibles biomarcadores integrando datos de metilación y expresión génica de alto rendimiento (152). Aunque identificaron 10 genes relacionados con el desarrollo del DG, el estudio carece de validación experimental o en una cohorte distinta. Wang et al. midieron sitios específicos de CpG, previamente publicados en la literatura, en 80 casos de DG y 80 controles en el primer trimestre como posibles predictores de la enfermedad. Un total de 13 sitios CpGs mostraron niveles de metilación del ADN estadísticamente significativos entre los casos de DG y los controles, aunque con un poder pronóstico medio-bajo (AUC varió de 0.593 a 0.650) (153). Recientemente, un estudio realizado por Megía et al. ha identificado varios sitios CpG como posibles biomarcadores para detectar el riesgo de desarrollo de anomalías en los niveles de glucosa más adelante en la vida en mujeres con DG (127). Los sitios CpG se asociaron con los genes *LINC00917*, *TRAPPC9* y *LEF1*.

En nuestro estudio identificamos tres sitios CpG en la cohorte EPIDG con una alta capacidad para discriminar entre mujeres embarazadas con DG y sin DG, de los cuales el cg04095097 se replicó en la muestra EPIPREG, lo que lo hace un excelente candidato. Una característica importante de nuestro estudio es que hemos encontrado marcas epigenéticas estables durante el embarazo capaces de diferenciar entre los grupos de DG y no DG, lo que nos hace pensar que estas marcas podrían estar presentes en el primer trimestre del embarazo. En la cohorte EPIDG, los casos de DG y

control se emparejaron por edad, edad gestacional e IMC pregestacional, mientras que en la cohorte de replicación no se parearon las muestras. Esto podría explicar por qué estos DMPs tenían menos capacidad para discriminar en la cohorte EPIPREG, al tratarse de una población más heterogénea. En la cohorte de descubrimiento, los casos de DG y los no DG eran muy homogéneos, lo que reduce la probabilidad de sesgo o variables de confusión que puedan influir en los resultados. Los sitios CpGs replicados parecen estar sólidamente asociados con la DG, ya que fueron comunes en las dos cohortes independientes que utilizan diferentes criterios de DG y de diferentes orígenes étnicos. Además, los mQTLs encontrados sugieren que estos CpGs están influenciados por variantes genéticas asociadas con la DG, rasgos cardio-metabólicos y enfermedades autoinmunes (154–156). No pudimos analizar el efecto de la metilación del ADN en la expresión génica debido a la falta de muestras de ARN.

Perfil epigenético diferencial de la descendencia de madres con y sin DG mantenido a lo largo del primer año de vida

En los últimos años ha habido un creciente interés en el estudio de marcas epigenéticas en saliva, como un biomarcador no invasivo con gran potencial. En relación con la DG algunos autores han realizado varias aproximaciones estudiando el patrón de metilación del ADN de niños nacidos de madres con DG y sin DG, pero hasta la fecha no se ha evaluado su evolución en el tiempo. Se han descrito perfiles epigenéticos relacionados con distintos fenotipos como el peso al nacer y estrés de la madre (157,158).

En este trabajo hemos visto, en primer lugar, que hay un perfil de metilación del ADN diferente en los niños nacidos de madres con DG y sin DG. Además, hemos observado que 16 de estos sitios CpGs se mantienen diferencialmente metilados a lo largo del tiempo, en la misma dirección. Algunos de estos CpGs se encuentran en los genes *FARP2* y *CYTH3* que están relacionados con rutas implicadas en procesos de inflamación y con DMT2. *FARP2* es un gen que permite la actividad del factor de intercambio de guanil-nucleótidos, también participa en la transducción de señales de la proteína Rac y la remodelación neuronal. Se localiza principalmente en el citoplasma (159). Mediante GWAS se ha visto que polimorfismos de este gen se encuentran relacionados con las concentraciones de HDL (160–162) y con la talla (168). Por otro lado, el gen *CYTH3* (cytohesin 3) codifica un miembro de la familia PSCD. La función principal de estos genes es el control de la estructura y función del aparato de Golgi, y puede tener función en la regulación del factor ADP-ribosilación, además de interactuar con ARF1. De acuerdo con algunos estudios de GWAS, se ha relacionado con el índice de masa corporal (163,164), y la talla (165).

En nuestro estudio, también observamos que 8 de los CpGs se relacionaron con variables antropométricas del niño. De estos, solo el cg10177795 se encontraba anotado en el gen *FARP2*. Este CpG correlacionaba con el percentil del perímetro cefálico al año, percentil de peso a los dos años y con el percentil de la talla a los dos años. El gen *FAMB5B* (también conocido como *BRINP2*) parece que se encuentra involucrado en la respuesta celular al ácido retinoico, regulación negativa del ciclo en células mitóticas y regulación positiva en la diferenciación de neuronas (166). De acuerdo con estudios de GWAS, este gen tiene SNPs relacionados con la circunferencia

de la cintura (167), síndrome metabólico (168), el índice de masa corporal (164,169,170), peso (169),talla (165,171) y obesidad (172). El cg10177795 se ha relacionado también con la edad (144). En nuestro estudio hemos visto que los niveles de metilación de este sitio CpG puede estar relacionado con los percentiles del perímetro cefálico y el peso tanto tras el parto como al año; así como con la talla. Otro CpG de interés es el cg17052441, el cual no está anotado a ningún gen, y hasta donde conocemos no hay trabajos en la literatura sobre el mismo. En nuestros resultados este sitio se relacionó con el perímetro cefálico, el peso y la talla en diferentes tiempos.

Por lo tanto, de este estudio podemos ver que los genes *FAM5B* y *FARP2* pueden ser de gran interés en el desarrollo del niño ya que son genes que han sido relacionados con índice de masa corporal, y con los niveles de HDL pudiendo esto ser un factor para el desarrollo de futuras enfermedades metabólicas. No obstante, el principal objetivo de este trabajo era comprobar si efectivamente, la DG, podía estar influyendo sobre el epigenoma de la descendencia, lo que hemos confirmado con nuestros resultados.

Con relación al riesgo de padecer obesidad, hemos encontrado que existe un perfil diferenciado de la metilación del ADN en algunos de los genes previamente asociados con la obesidad infantil en la literatura.

En conclusión, vemos como existe un perfil epigenético que diferencia a los niños nacidos de madres con DG de aquellos de madres sin DG, y algunas de estas marcas se mantienen en el tiempo. Además, observamos que algunos CpGs se correlacionan con variables antropométricas en los dos

momentos, sugiriendo que pueden ser de gran interés en el desarrollo de problemas metabólicos en el futuro. Así el análisis epigenético en muestras de saliva puede ser de utilidad para la predicción del desarrollo de alteraciones en el patrón de crecimiento en los niños.

Asociación de un MRS de DMT2 con la DG

Este es el primer estudio que demuestra que un MRS para la diabetes tipo 2 (DMT2) se asocia con la diabetes gestacional (DG), en dos cohortes con criterios de diagnóstico y ancestros diferentes. Además, el MRS para la DMT2 se asoció con niveles más altos de glucosa en ayunas. Seis sitios CpG de los 42 incluidos en el MRS se asociaron significativamente con la DG en la cohorte europea de EPIPREG, sudasiáticos de EPIPREG y mediterráneos de EPIDG. Identificamos 23 mQTLs (locus de caracteres cuantitativos relacionados con la metilación) vinculados a los 6 sitios CpG, que se asociaron con fenotipos como la DMT2 y DG entre otros.

En los últimos años numerosos trabajos se han enfocado en los GRS, los cuales permiten cuantificar el riesgo acumulado de una enfermedad causado por variantes genéticas con tamaños de efectos pequeños en múltiples genes identificados en los estudios de GWAS. A diferencia de los GRS, el MRS podría reflejar el efecto de los factores ambientales sobre la enfermedad o el fenotipo y ayuda a comprender el mecanismo de la enfermedad.

LA DG y la DMT2 comparten muchos factores fisiopatológicos, pero los mecanismos subyacentes exactos son en su mayoría desconocidos. Existe la necesidad de comprender cómo se relacionan la DMT2 y la DG para mejorar la prevención y evitar complicaciones metabólicas en el futuro. Se han utilizado los estudios de GWAS para investigar el posible vínculo entre

la DG y la DMT2 (57), pero hasta ahora ninguno ha probado si comparten marcas epigenéticas. Un número creciente de estudios sugiere que los factores ambientales están asociados con marcas epigenéticas y que esas marcas pueden ser importantes para comprender las interacciones entre la genética y los factores ambientales. Hasta ahora, se han publicado algunos EWAS de DG en sangre periférica materna, pero los datos no son suficientes para crear un MRS para la DG debido a la heterogeneidad en los criterios diagnósticos utilizados, así como al tamaño muestral reducido (127,128,173,174).

Nuestro MRS se construyó a partir de un metaanálisis de cinco cohortes europeas que identificaron sitios CpG asociados con la incidencia de la DMT2 (119). Algunos de los genes incluidos en el MRS están asociados con el metabolismo de la glucosa (175,176), la DMT2, la diabetes mellitus tipo 1 y la DG (177). ATP binding cassette subfamily G member 1 (*ABCG1*) es un gen involucrado en el transporte de colesterol y fosfolípidos y puede regular la homeostasis lipídica celular. *ABCG1* se ha asociado con la diabetes y el síndrome metabólico (178). Además, algunos estudios muestran que la DG afecta la homeostasis del colesterol a través de este gen, por lo que podría desempeñar un papel en la DG y los eventos cardiovasculares (179).

Además, los resultados que hemos obtenido reflejan como el ambiente geográfico puede influir también en los niveles de metilación, incluso más que el origen étnico, como muestran los diferentes niveles de MRS en la cohorte EPIDG respecto a EPIPREG. El MRS de EPIPREG es muy parecido entre las mujeres surasiáticas y las europeas, mientras que difiere de la cohorte EPIDG.

Según los resultados de mQTL, las marcas epigenéticas relacionadas con la DMT2 en los genes *MAP4K2*, *SLC9A3R1*, *LGAL3BP*, *SYNGR1*, *ABCG1* y *SLC9A1* también se asocian con un mayor riesgo de DG. *MAP4K2* codifica una proteína miembro de la familia de quinasas de serina/treonina y participa en la diferenciación de las células β (180). *SLC9A3R1*, también conocida como proteína de andamiaje PDZ de la familia NHERF 1 (*NHERF*), codifica un cofactor regulador del transportador de sodio/hidrógeno. Es un componente clave de las vías de señalización de la insulina y el IGF-1 y se ha relacionado con marcas epigenéticas de la DMT2 (175). Un EWAS realizado en una cohorte sudafricana sugirió que las marcas epigenéticas en *SLC9A3* están asociadas con la DG (128).

Además, las variantes genéticas relacionadas con cg06378491 y cg22650271 (dentro de los genes *MAP4K2* y *SYNGR1*) se asociaron con la DG con significación nominal. El rs7289325, el mQTL de cg22650271, se encuentra en el factor de crecimiento derivado de plaquetas subunidad B (*PDGFB*). Este gen codifica un miembro de la familia de proteínas compuesta tanto por el factor de crecimiento derivado de plaquetas (*PDGF*) como por el factor de crecimiento endotelial vascular. Estas proteínas se unen y activan los receptores tirosina quinasa del receptor *PDGF*, que desempeñan un papel en una amplia gama de procesos del desarrollo (181). Esta familia de proteínas se ha relacionado con una disminución de la función de las células beta y el nivel de proteína *PDGFB* aumenta en los pacientes con DG y también reduce la función de las células beta durante la gestación a través de una peor diferenciación de las células beta (182).

La combinación de dos cohortes independientes, dos ancestros en EPIPREG y el uso de diferentes criterios diagnósticos de DG, refuerza nuestros resultados. El hecho de que el MRS funcione en ambas cohortes indica que los criterios utilizados son suficientes para ver la relación entre la DG y la DMT2 a través de la epigenética. Además, nuestros resultados muestran que el riesgo epigenético de la DMT2 también está relacionado con la DG, lo que sugiere que muchas marcas de metilación pueden ser comunes en ambas patologías. La prevalencia y estudios de DG son mucho menores que los que hay de DM2, por lo que se dispone de un menor tamaño muestral para este tipo de investigaciones. En nuestro caso, el número de participantes también puede ser un factor limitante, pues resulta en una menor potencia estadística para probar dicha asociación. Seguramente, un MRS para la DG basado en EWAS de la DG podría haber proporcionado una predicción más precisa. Sin embargo, los EWAS disponibles de la DG en leucocitos de sangre periférica materna u otros tejidos maternos son pocos, los tamaños de muestra son bastante pequeños, hay menos diversidad ancestral y los diferentes criterios de diagnóstico complican la comparación entre cohortes. Son escasos los estudios con tamaños de muestra grandes (183,184). Finalmente, nuestro diseño de estudio transversal no puede determinar si el estado de DG influye en la metilación del ADN o viceversa.

Este estudio mostró que las marcas epigenéticas asociadas con la DMT2 también se asociaron con la DG. Las investigaciones futuras deberían comparar si el GRS y el MRS o una combinación de ambos proporcionan mejores predicciones de DG y DMT2 (10). En resumen, hemos mostrado por primera vez que el MRS para la DMT2 está significativamente asociado

con la DG, lo que sugiere mecanismos epigenéticos compartidos entre la DG y la T2DM. Esto puede ayudar a explicar algunos de los mecanismos moleculares que median en el aumento del riesgo de desarrollar DMT2 después de haber tenido DG. Además, comprender qué factores modifican la metilación de estos genes y el papel de las variantes genéticas en el desarrollo de la enfermedad puede ayudar a mejorar la prevención y el manejo tanto de la DG como de la DMT2 en el futuro.

Comentario final

Nuestros resultados apuntan a que el ambiente adverso causado por la DG tiene un efecto sobre el epigenoma de la madre y su descendencia, a través de cambios en la metilación del ADN en genes implicados en rutas metabólicas relacionadas con la DG tales como la señalización de la insulina, respuesta a los niveles de nutrientes, resistencia a la insulina, entre otros. Asimismo, muchas de estas marcas se mantienen estables en el tiempo, y podrían servir como biomarcadores de riesgo de futura diabetes u otras alteraciones metabólicas relacionadas, tanto en las madres como en su descendencia.

Como cualquier estudio de investigación, esta tesis presenta una serie de limitaciones, que, aunque algunas han sido ya comentadas, nos gustaría resumir como comentario final y futuras aproximaciones. Se ha realizado un análisis exploratorio en un subgrupo de pacientes de la cohorte EPIDG, y algunos de sus resultados deben de ser validados bien en la cohorte completa o en otras cohortes de mayor tamaño muestral, para corroborar su validez. Asimismo, el uso de muestras de sangre periférica y saliva conlleva limitaciones a la hora de explicar su relevancia en la fisiopatología del fenotipo, siendo su principal ventaja su accesibilidad como posibles

biomarcadores clínicos. Este trabajo ha generado hipótesis que necesitan ser confirmadas con más estudios en el futuro. Por último, el ajuste estadístico por variables como el IMC o el HOMA-IR puede que impidan la detección de DMPs que funciones precisamente a través de mecanismos como la obesidad o la resistencia a la insulina.

Conclusiones

Con todos estos resultados, las conclusiones han sido las siguientes:

- 1- Existe un patrón de metilación del ADN diferente entre mujeres embarazadas con DG y sin DG.
- 2- Los genes anotados en los sitios diferencialmente metilados (DMPs) podrían estar involucrados en el desarrollo de la DG y en alteraciones en variables metabólicas relacionadas.
- 3- Hay marcas epigenéticas, basadas en la metilación del ADN, que se mantienen diferencialmente metiladas a lo largo del embarazo, entre mujeres con DG y sin DG.
- 4- Hemos identificado sitios CpGs específicos con una gran capacidad de discriminar entre embarazadas con DG y sin DG, lo que podría convertirlos en candidatos para biomarcadores de predicción de DG en el futuro.
- 5- Encontramos un patrón de metilación del ADN diferenciado en las muestras de saliva entre los niños nacidos de madres que padecieron DG en comparación con los niños nacidos de madres sin DG.
- 6- Algunas de estas marcas de metilación del ADN persisten a la edad de un año y se asocian con variables relacionadas con el patrón de crecimiento del niño.
- 7- Por primera vez, mostramos que las marcas de metilación del ADN para T2DM también se asocian con la DG, lo que sugiere mecanismos epigenéticos compartidos entre ambas patologías.

Conclusions

With all these results, the conclusions are as follows:

- 1- There is a different DNA methylation pattern between pregnant women with GD and those without GD.
- 2- Genes annotated at differentially methylated sites (DMPs) could be involved in the development of GD and alterations in related metabolic variables.
- 3- There are epigenetic marks, based on DNA methylation, that remain differentially methylated throughout pregnancy, between women with and without GD.
- 4- We have identified specific CpGs sites with a strong ability to discriminate between pregnant women with and without GD, that could be potential predictive biomarkers.
- 5- We found a distinct DNA methylation pattern in saliva samples among children born to mothers with GD compared to children born to mothers without GD.
- 6- Some of these DNA methylation marks persist at the age of one year and are associated with variables related to the child's growth pattern.
- 7- For the first time, we show that DNA methylation marks for T2DM are also associated with GDM, suggesting shared epigenetic mechanisms between both pathologies.

Bibliografía

1. WHO European Regional Obesity Report 2022 [Internet]. [cited 2023 May 30]. Available from: https://apps.who.int/iris/handle/10665/353747?search-result=true&query=WHO+EUROPEAN+REGIONAL+OBESITY+REPORT+2022&scope=&rpp=10&sort_by=score&order=desc
2. Boutari C, Mantzoros CS. A 2022 update on the epidemiology of obesity and a call to action: as its twin COVID-19 pandemic appears to be receding, the obesity and dysmetabolism pandemic continues to rage on. *Metabolism* [Internet]. 2022 Aug 1 [cited 2023 May 30];133:155217. Available from: </pmc/articles/PMC9107388/>
3. Chatterjee S, Khunti K, Davies MJ. Type 2 diabetes. *Lancet* [Internet]. 2017 Jun 3 [cited 2023 May 18];389(10085):2239–51. Available from: <http://www.thelancet.com/article/S0140673617300582/fulltext>
4. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Kaabi J Al. Epidemiology of Type 2 Diabetes – Global Burden of Disease and Forecasted Trends. *J Epidemiol Glob Health* [Internet]. 2020 Mar 1 [cited 2023 May 30];10(1):107. Available from: </pmc/articles/PMC7310804/>
5. Heindel JJ, Balbus J, Birnbaum L, Brune-Drisse MN, Grandjean P, Gray K, et al. Developmental origins of health and disease: Integrating environmental influences [Internet]. Vol. 156, *Endocrinology*. Endocrine Society; 2015 [cited 2020 Sep 21]. p. 3416–21. Available from: <https://pubmed.ncbi.nlm.nih.gov/26241070/>
6. Moore TG, Arefadib N, Deery A, West S. The first thousand days: an evidence paper. 2017.
7. Damm P, Houshmand-Oeregaard A, Kelstrup L, Lauenborg J, Mathiesen ER, Clausen TD. Gestational diabetes mellitus and long-term consequences for mother and offspring: a view from Denmark [Internet]. Vol. 59, *Diabetologia*. Springer Verlag; 2016 [cited 2020 Sep 28]. p. 1396–9. Available from: <https://link.springer.com/article/10.1007/s00125-016-3985-5>
8. DABELEA D, HANSON RL, BENNETT PH, PETTITT DJ, LINDSAY RS,

- IMPERATORE G, et al. Intrauterine Exposure to Diabetes Conveys Risk for Diabetes and Obesity in Offspring Above That Attributable to Genetics. *Diabetes* [Internet]. 1999 May 1 [cited 2020 Sep 28];48(5):SA52–SA52. Available from: <https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=00121797&v=2.1&it=r&id=GALE%7CA54770543&sid=googleScholar&linkaccess=fulltext>
9. Association AD. Gestational Diabetes Mellitus. *Diabetes Care* [Internet]. 2003 Jan 1 [cited 2023 Jun 5];26(suppl_1):s103–5. Available from: https://diabetesjournals.org/care/article/26/suppl_1/s103/21729/Gestational-Diabetes-Mellitus
 10. Elsayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, et al. 15. Management of Diabetes in Pregnancy: Standards of Care in Diabetes—2023. *Diabetes Care* [Internet]. 2023 Jan 1 [cited 2023 Mar 29];46(Supplement_1):S254–66. Available from: https://diabetesjournals.org/care/article/46/Supplement_1/S254/148052/15-Management-of-Diabetes-in-Pregnancy-Standards
 11. Lesseur C, Chen J. Adverse Maternal Metabolic Intrauterine Environment and Placental Epigenetics: Implications for Fetal Metabolic Programming. *Curr Environ Heal reports*. 2018 Dec 1;5(4):531–43.
 12. Zhu Y, Zhang C. Prevalence of Gestational Diabetes and Risk of Progression to Type 2 Diabetes: a Global Perspective. *Curr Diab Rep* [Internet]. 2016 Jan 1 [cited 2023 Jun 26];16(1):7. Available from: </pmc/articles/PMC6675405/>
 13. Wang H, Li N, Chivese T, Werfalli M, Sun H, Yuen L, et al. IDF Diabetes Atlas: Estimation of Global and Regional Gestational Diabetes Mellitus Prevalence for 2021 by International Association of Diabetes in Pregnancy Study Group’s Criteria. *Diabetes Res Clin Pract* [Internet]. 2022 Jan 1 [cited 2023 Jun 14];183:4. Available from: <http://www.diabetesresearchclinicalpractice.com/article/S0168827221004095/fulltext>
 14. Corcoy R, Lumbreras B, Bartha JL, Ricart W. Nuevos criterios

diagnósticos de diabetes mellitus gestacional a partir del estudio HAPO. ¿Son válidos en nuestro medio? *Endocrinol y Nutr* [Internet]. 2010 Jun 1 [cited 2022 Mar 2];57(6):277–80. Available from: <https://www.elsevier.es/es-revista-endocrinologia-nutricion-12-articulo-nuevos-criterios-diagnosticos-diabetes-mellitus-S157509221000104X>

15. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. *Diabetes* [Internet]. 1979 [cited 2022 Jun 10];28(12):1039–57. Available from: <https://pubmed.ncbi.nlm.nih.gov/510803/>
16. Carpenter MW, Coustan DR. Criteria for screening tests for gestational diabetes. *Am J Obstet Gynecol* [Internet]. 1982 Dec 1 [cited 2023 Aug 1];144(7):768–73. Available from: <https://pubmed.ncbi.nlm.nih.gov/7148898/>
17. Association AD. Standards of Medical Care in Diabetes—2011. *Diabetes Care* [Internet]. 2011 Jan 1 [cited 2023 Aug 1];34(Supplement_1):S11–61. Available from: <https://dx.doi.org/10.2337/dc11-S011>
18. Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy: a World Health Organization Guideline. *Diabetes Res Clin Pract* [Internet]. 2014 [cited 2023 Feb 27];103(3):341–63. Available from: <https://pubmed.ncbi.nlm.nih.gov/24847517/>
19. Fakhrol-Alam M, Sharmin-Jahan, Mashfiqul-Hasan, Nusrat-Sultana, Mohona-Zaman, Rakibul-Hasan M, et al. Insulin secretory defect may be the major determinant of GDM in lean mothers. *J Clin Transl Endocrinol* [Internet]. 2020 Jun 1 [cited 2022 Aug 9];20. Available from: <https://pubmed.ncbi.nlm.nih.gov/32382513/>
20. Catalano PM, Tyzbir ED, Wolfe RR, Roman NM, B. Amini S, Sims EAH. Longitudinal changes in basal hepatic glucose production and suppression during insulin infusion in normal pregnant women. *Am J Obstet Gynecol* [Internet]. 1992 [cited 2023 Jun 30];167(4 Pt 1):913–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/1415425/>

21. Catalano PM, Tyzbir ED, Roman NM, Amini SB, Sims EAH. Longitudinal changes in insulin release and insulin resistance in nonobese pregnant women. *Am J Obstet Gynecol* [Internet]. 1991 [cited 2023 Jun 30];165(6 Pt 1):1667–72. Available from: <https://pubmed.ncbi.nlm.nih.gov/1750458/>
22. DeFronzo RA, Ferrannini E, Groop L, Henry RR, Herman WH, Holst JJ, et al. Type 2 diabetes mellitus. *Nat Rev Dis Prim* 2015 11 [Internet]. 2015 Jul 23 [cited 2023 Jul 5];1(1):1–22. Available from: <https://www.nature.com/articles/nrdp201519>
23. Barbour LA, McCurdy CE, Hernandez TL, Kirwan JP, Catalano PM, Friedman JE. Cellular Mechanisms for Insulin Resistance in Normal Pregnancy and Gestational Diabetes. *Diabetes Care* [Internet]. 2007 Jul 1 [cited 2023 Jul 5];30(Supplement_2):S112–9. Available from: <https://dx.doi.org/10.2337/dc07-s202>
24. Friedman JE, Ishizuka T, Shao J, Huston L, Highman T, Catalano P. Impaired glucose transport and insulin receptor tyrosine phosphorylation in skeletal muscle from obese women with gestational diabetes. *Diabetes* [Internet]. 1999 [cited 2023 Jul 5];48(9):1807–14. Available from: <https://pubmed.ncbi.nlm.nih.gov/10480612/>
25. Catalano PM. Trying to understand gestational diabetes. *Diabet Med* [Internet]. 2014 Mar [cited 2023 Jul 7];31(3):273–81. Available from: <https://pubmed.ncbi.nlm.nih.gov/24341419/>
26. Friedman JE, Kirwan JP, Jing M, Presley L, Catalano PM. Increased skeletal muscle tumor necrosis factor-alpha and impaired insulin signaling persist in obese women with gestational diabetes mellitus 1 year postpartum. *Diabetes* [Internet]. 2008 Mar [cited 2023 Jul 7];57(3):606–13. Available from: <https://pubmed.ncbi.nlm.nih.gov/18083784/>
27. David McIntyre H. Discovery, Knowledge, and Action-Diabetes in Pregnancy Across the Translational Spectrum: The 2016 Norbert Freinkel Award Lecture. *Diabetes Care* [Internet]. 2018 Feb 1 [cited 2023 Jul 25];41(2):227–32. Available from: <https://pubmed.ncbi.nlm.nih.gov/29358466/>

28. Hotamisligil GS, Peraldi P, Budavari A, Ellis R, White MF, Spiegelman BM. IRS-1-mediated inhibition of insulin receptor tyrosine kinase activity in TNF- α - and obesity-induced insulin resistance. *Science* [Internet]. 1996 Feb 2 [cited 2023 Jul 25];271(5249):665–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/8571133/>
29. Kirwan JP, Varastehpour A, Jing M, Presley L, Shao J, Friedman JE, et al. Reversal of insulin resistance postpartum is linked to enhanced skeletal muscle insulin signaling. *J Clin Endocrinol Metab* [Internet]. 2004 Sep [cited 2023 Jul 25];89(9):4678–84. Available from: <https://pubmed.ncbi.nlm.nih.gov/15356080/>
30. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. *Nat Rev Dis Prim* 2019 51 [Internet]. 2019 Jul 11 [cited 2023 Jul 31];5(1):1–19. Available from: <https://www.nature.com/articles/s41572-019-0098-8>
31. Metzger BE, Lowe LP, Dyer AR, Trimble ER, Sheridan B, Hod M, et al. Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study: Associations With Neonatal Anthropometrics. *Diabetes* [Internet]. 2009 Feb [cited 2023 Jul 31];58(2):453. Available from: </pmc/articles/PMC2628620/>
32. Crowther CA, Hiller JE, Moss JR, McPhee AJ, Jeffries WS, Robinson JS. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. *N Engl J Med* [Internet]. 2005 Jun 16 [cited 2023 Jul 31];352(24):2477–86. Available from: <https://pubmed.ncbi.nlm.nih.gov/15951574/>
33. Bellamy L, Casas JP, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. *Lancet (London, England)* [Internet]. 2009 [cited 2022 Dec 19];373(9677):1773–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/19465232/>
34. Lauenborg J, Hansen T, Jensen DM, Vestergaard H, Mølsted-Pedersen L, Hornnes P, et al. Increasing incidence of diabetes after gestational diabetes: a long-term follow-up in a Danish population. *Diabetes Care* [Internet]. 2004 May [cited 2023 Jul 31];27(5):1194–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/15111544/>

35. Kramer CK, Campbell S, Retnakaran R. Gestational diabetes and the risk of cardiovascular disease in women: a systematic review and meta-analysis. *Diabetologia* [Internet]. 2019 Jun 1 [cited 2023 Oct 21];62(6):905–14. Available from: <https://pubmed.ncbi.nlm.nih.gov/30843102/>
36. Lowe WL, Scholtens DM, Lowe LP, Kuang A, Nodzenski M, Talbot O, et al. Association of Gestational Diabetes With Maternal Disorders of Glucose Metabolism and Childhood Adiposity. *JAMA*. 2018 Sep;320(10):1005–16.
37. Clausen TD, Mathiesen ER, Hansen T, Pedersen O, Jensen DM, Lauenborg J, et al. High prevalence of type 2 diabetes and pre-diabetes in adult offspring of women with gestational diabetes mellitus or type 1 diabetes: the role of intrauterine hyperglycemia. *Diabetes Care* [Internet]. 2008 Feb [cited 2023 Jul 31];31(2):340–6. Available from: <https://pubmed.ncbi.nlm.nih.gov/18000174/>
38. Clausen TD, Mathiesen ER, Hansen T, Pedersen O, Jensen DM, Lauenborg J, et al. Overweight and the metabolic syndrome in adult offspring of women with diet-treated gestational diabetes mellitus or type 1 diabetes. *J Clin Endocrinol Metab* [Internet]. 2009 [cited 2023 Jul 31];94(7):2464–70. Available from: <https://pubmed.ncbi.nlm.nih.gov/19417040/>
39. Kelstrup L, Damm P, Mathiesen ER, Hansen T, Vaag AA, Pedersen O, et al. Insulin resistance and impaired pancreatic β -cell function in adult offspring of women with diabetes in pregnancy. *J Clin Endocrinol Metab* [Internet]. 2013 Sep [cited 2023 Jul 31];98(9):3793–801. Available from: <https://pubmed.ncbi.nlm.nih.gov/23796568/>
40. Lowe WL, Scholtens DM, Kuang A, Linder B, Lawrence JM, Lebenthal Y, et al. Hyperglycemia and Adverse Pregnancy Outcome Follow-up Study (HAPO FUS): Maternal Gestational Diabetes Mellitus and Childhood Glucose Metabolism. *Diabetes Care* [Internet]. 2019 Mar 1 [cited 2023 Jul 31];42(3):372–80. Available from: <https://dx.doi.org/10.2337/dc18-1646>
41. Zhang C, Rawal S, Chong YS. Risk factors for gestational diabetes: is prevention possible? *Diabetologia* [Internet]. 2016 Jul 1 [cited 2023

- Jun 22];59(7):1385–90. Available from: <https://pubmed.ncbi.nlm.nih.gov/27165093/>
42. Torloni MR, Betrán AP, Horta BL, Nakamura MU, Atallah AN, Moron AF, et al. Prepregnancy BMI and the risk of gestational diabetes: a systematic review of the literature with meta-analysis. *Obes Rev* [Internet]. 2009 Mar [cited 2023 Jun 27];10(2):194–203. Available from: <https://pubmed.ncbi.nlm.nih.gov/19055539/>
 43. Bao W, Michels KB, Tobias DK, Li S, Chavarro JE, Gaskins AJ, et al. Parental smoking during pregnancy and the risk of gestational diabetes in the daughter. *Int J Epidemiol* [Internet]. 2016 Feb 1 [cited 2023 Jun 27];45(1):160–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/26748845/>
 44. Phelan S, Jelalian E, Coustan D, Caughey AB, Castorino K, Hagobian T, et al. Protocol for a randomized controlled trial of pre-pregnancy lifestyle intervention to reduce recurrence of gestational diabetes: Gestational Diabetes Prevention/Prevención de la Diabetes Gestacional. *Trials* [Internet]. 2021 Dec 1 [cited 2022 Feb 2];22(1). Available from: </pmc/articles/PMC8024941/>
 45. Zhang C, Solomon CG, Manson JAE, Hu FB. A prospective study of pregravid physical activity and sedentary behaviors in relation to the risk for gestational diabetes mellitus. *Arch Intern Med*. 2006 Mar 13;166(5):543–8.
 46. Zhang C, Qiu C, Hu FB, David RM, van Dam RM, Bralley A, et al. Maternal Plasma 25-Hydroxyvitamin D Concentrations and the Risk for Gestational Diabetes Mellitus. *PLoS One* [Internet]. 2008 Nov 18 [cited 2023 Jun 27];3(11). Available from: </pmc/articles/PMC2582131/>
 47. Zhang C, Williams MA, Sorensen TK, King IB, Kestin MM, Thompson M Lou, et al. Maternal plasma ascorbic acid (vitamin C) and risk of gestational diabetes mellitus. *Epidemiology* [Internet]. 2004 Sep [cited 2023 Jun 27];15(5):597–604. Available from: https://journals.lww.com/epidem/Fulltext/2004/09000/Maternal_Plasma_Ascorbic_Acid__Vitamin_C__and_Risk.18.aspx
 48. Dominguez LJ, Martínez-González MA, Basterra-Gortari FJ, Gea A,

- Barbagallo M, Bes-Rastrollo M. Fast Food Consumption and Gestational Diabetes Incidence in the SUN Project. *PLoS One* [Internet]. 2014 Sep 1 [cited 2023 Jun 27];9(9). Available from: </pmc/articles/PMC4162567/>
49. Retnakaran R, Kramer CK, Ye C, Kew S, Hanley AJ, Connelly PW, et al. Fetal sex and maternal risk of gestational diabetes mellitus: the impact of having a boy. *Diabetes Care* [Internet]. 2015 May 1 [cited 2023 Jun 30];38(5):844–51. Available from: <https://pubmed.ncbi.nlm.nih.gov/25693837/>
 50. Rauh-Hain JA, Rana S, Tamez H, Wang A, Cohen B, Cohen A, et al. Risk for developing gestational diabetes in women with twin pregnancies. *J Matern Fetal Neonatal Med* [Internet]. 2009 [cited 2023 Jun 30];22(4):293–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/19340713/>
 51. Hedderson M, Ehrlich S, Sridhar S, Darbinian J, Moore S, Ferrara A. Racial/Ethnic Disparities in the Prevalence of Gestational Diabetes Mellitus by BMI. *Diabetes Care* [Internet]. 2012 Jul [cited 2023 Jun 30];35(7):1492. Available from: </pmc/articles/PMC3379591/>
 52. Flannick J, Florez JC. Type 2 diabetes: Genetic data sharing to advance complex disease research [Internet]. Vol. 17, *Nature Reviews Genetics*. Nature Publishing Group; 2016 [cited 2020 Sep 28]. p. 535–49. Available from: <https://www.nature.com/articles/nrg.2016.56>
 53. Ellard S, Bellanné-Chantelot C, Hattersley AT, Carette C, Castano Gonzalez L, De Nanclares Leal G, et al. Best practice guidelines for the molecular genetic diagnosis of maturity-onset diabetes of the young. *Diabetologia* [Internet]. 2008 Apr [cited 2023 Jul 25];51(4):546. Available from: </pmc/articles/PMC2270360/>
 54. Wu NN, Zhao D, Ma W, Lang JN, Liu SM, Fu Y, et al. A genome-wide association study of gestational diabetes mellitus in Chinese women. *J Matern Fetal Neonatal Med* [Internet]. 2021 [cited 2023 Jun 22];34(10):1557–64. Available from: <https://pubmed.ncbi.nlm.nih.gov/31269844/>
 55. Kwak SH, Kim SH, Cho YM, Go MJ, Cho YS, Choi SH, et al. A genome-

- wide association study of gestational diabetes mellitus in Korean women. *Diabetes* [Internet]. 2012 Feb [cited 2023 Jun 22];61(2):531–41. Available from: <https://pubmed.ncbi.nlm.nih.gov/22233651/>
56. Park S. Interaction of Polygenetic Variants for Gestational Diabetes Mellitus Risk with Breastfeeding and Korean Balanced Diet to Influence Type 2 Diabetes Risk in Later Life in a Large Hospital-Based Cohort. *J Pers Med* [Internet]. 2021 Nov 1 [cited 2023 Jun 22];11(11). Available from: <https://pubmed.ncbi.nlm.nih.gov/34834527/>
 57. Pervjakova N, Moen GH, Borges MC, Ferreira T, Cook JP, Allard C, et al. Multi-ancestry genome-wide association study of gestational diabetes mellitus highlights genetic links with type 2 diabetes. *Hum Mol Genet* [Internet]. 2022 Sep 29 [cited 2023 May 15];31(19):3377–91. Available from: <https://pubmed.ncbi.nlm.nih.gov/35220425/>
 58. Dias S, Pfeiffer C, Abrahams Y, Rheeder P, Adam S. Molecular biomarkers for gestational diabetes mellitus. *Int J Mol Sci*. 2018 Oct 1;19(10).
 59. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature*. 2004 May 27;429(6990):457–63.
 60. Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond [Internet]. Vol. 13, *Nature Reviews Genetics*. Nat Rev Genet; 2012 [cited 2020 Sep 28]. p. 484–92. Available from: <https://pubmed.ncbi.nlm.nih.gov/22641018/>
 61. Visone R, Bacalini MG, Franco S Di, Ferracin M, Colorito ML, Pagotto S, et al. DNA methylation of shelf, shore and open sea CpG positions distinguish high microsatellite instability from low or stable microsatellite status colon cancer stem cells. *Epigenomics* [Internet]. 2019 May 8 [cited 2023 Nov 14];11(6):587–604. Available from: <https://www.futuremedicine.com/doi/10.2217/epi-2018-0153>
 62. Robertson KD, Wolffe AP. DNA methylation in health and disease.

- Nat Rev Genet 2000 11 [Internet]. 2000 [cited 2023 Aug 3];1(1):11–9. Available from: <https://www.nature.com/articles/35049533>
63. Bird A. Perceptions of epigenetics. Nat 2007 4477143 [Internet]. 2007 May 23 [cited 2023 Aug 3];447(7143):396–8. Available from: <https://www.nature.com/articles/nature05913>
 64. Kouzarides T. Chromatin modifications and their function. Cell [Internet]. 2007 Feb 23 [cited 2023 Aug 9];128(4):693–705. Available from: <https://pubmed.ncbi.nlm.nih.gov/17320507/>
 65. Berger SL. The complex language of chromatin regulation during transcription. Nat 2007 4477143 [Internet]. 2007 May 23 [cited 2023 Aug 9];447(7143):407–12. Available from: <https://www.nature.com/articles/nature05915>
 66. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res [Internet]. 2011 Mar [cited 2023 Aug 9];21(3):381–95. Available from: <https://pubmed.ncbi.nlm.nih.gov/21321607/>
 67. Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nat Rev Genet [Internet]. 2014 [cited 2023 Aug 9];15(6):394–408. Available from: <https://pubmed.ncbi.nlm.nih.gov/24805120/>
 68. Wei JW, Huang K, Yang C, Kang CS. Non-coding RNAs as regulators in epigenetics (Review). Oncol Rep [Internet]. 2017 Jan 1 [cited 2023 Oct 1];37(1):3–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/27841002/>
 69. Ruvkun G. Molecular biology. Glimpses of a tiny RNA world. Science [Internet]. 2001 Oct 26 [cited 2023 Oct 1];294(5543):797–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/11679654/>
 70. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell [Internet]. 2009 Jan 23 [cited 2023 Oct 1];136(2):215–33. Available from: <https://pubmed.ncbi.nlm.nih.gov/19167326/>
 71. Moazed D. Small RNAs in transcriptional gene silencing and genome defence. Nature [Internet]. 2009 Jan 22 [cited 2023 Oct

- 1];457(7228):413–20. Available from:
<https://pubmed.ncbi.nlm.nih.gov/19158787/>
72. Jinek M, Doudna JA. A three-dimensional view of the molecular machinery of RNA interference. *Nature* [Internet]. 2009 Jan 22 [cited 2023 Oct 1];457(7228):405–12. Available from:
<https://pubmed.ncbi.nlm.nih.gov/19158786/>
73. Grivna ST, Beyret E, Wang Z, Lin H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev* [Internet]. 2006 Jul 1 [cited 2023 Oct 1];20(13):1709–14. Available from:
<https://pubmed.ncbi.nlm.nih.gov/16766680/>
74. Lin H. piRNAs in the germ line. *Science* [Internet]. 2007 Apr 20 [cited 2023 Oct 1];316(5823):397. Available from:
<https://pubmed.ncbi.nlm.nih.gov/17446387/>
75. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet* [Internet]. 2009 Mar [cited 2023 Oct 1];10(3):155–9. Available from:
<https://pubmed.ncbi.nlm.nih.gov/19188922/>
76. Flanagan JM, Brook MN, Orr N, Tomczyk K, Coulson P, Fletcher O, et al. Temporal stability and determinants of white blood cell DNA methylation in the breakthrough generations study. *Cancer Epidemiol Biomarkers Prev* [Internet]. 2015 Jan 1 [cited 2023 Aug 10];24(1):221–9. Available from:
<https://pubmed.ncbi.nlm.nih.gov/25371448/>
77. Maddocks ODK, Labuschagne CF, Adams PD, Vousden KH. Serine Metabolism Supports the Methionine Cycle and DNA/RNA Methylation through De Novo ATP Synthesis in Cancer Cells. *Mol Cell* [Internet]. 2016 Jan 1 [cited 2023 Aug 10];61(2):210. Available from: /pmc/articles/PMC4728077/
78. Grunau C, Clark SJ, Rosenthal A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res* [Internet]. 2001 Jul 1 [cited 2023 Aug 14];29(13):65–65. Available from:
<https://pubmed.ncbi.nlm.nih.gov/11433041/>

79. Derks S, Lentjes MHFM, Hellebrekers DMEI, De Bruïne AP, Herman JG, Van Engeland M. Methylation-specific PCR unraveled. *Cell Oncol* [Internet]. 2004 [cited 2023 Aug 14];26(5–6):291–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/15623939/>
80. Poulin M, Zhou JY, Yan L, Shioda T. Pyrosequencing Methylation Analysis. In 2018 [cited 2020 Apr 23]. p. 283–96. Available from: http://link.springer.com/10.1007/978-1-4939-8751-1_17
81. Bibikova M, Le J, Barnes B, Saedinia-Melnyk S, Zhou L, Shen R, et al. Genome-wide DNA methylation profiling using Infinium® assay. *Epigenomics* [Internet]. 2009 Oct [cited 2023 Aug 10];1(1):177–200. Available from: <https://pubmed.ncbi.nlm.nih.gov/22122642/>
82. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, V Lord R, et al. De novo identification of differentially methylated regions in the human genome. *Epigenetics and Chromatin* [Internet]. 2015 Jan 27 [cited 2020 Nov 17];8(1):6. Available from: <https://epigeneticsandchromatin.biomedcentral.com/articles/10.1186/1756-8935-8-6>
83. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* [Internet]. 2012 May 8 [cited 2022 Jun 13];13(1):1–16. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-13-86>
84. Yong WS, Hsu FM, Chen PY. Profiling genome-wide DNA methylation. *Epigenetics Chromatin* [Internet]. 2016 Jun 29 [cited 2023 Oct 26];9(1). Available from: </pmc/articles/PMC4926291/>
85. Kingan SB, Heaton H, Cudini J, Lambert CC, Baybayan P, Galvin BD, et al. A High-Quality De novo Genome Assembly from a Single Mosquito Using PacBio Sequencing. *Genes (Basel)* [Internet]. 2019 Jan 1 [cited 2023 Oct 26];10(1). Available from: </pmc/articles/PMC6357164/>
86. Laszlo AH, Derrington IM, Brinkerhoff H, Langford KW, Nova IC, Samson JM, et al. Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proc Natl Acad Sci*

- U S A [Internet]. 2013 Nov 19 [cited 2023 Oct 26];110(47):18904–9. Available from: /pmc/articles/PMC3839702/
87. Campagna MP, Xavier A, Lechner-Scott J, Maltby V, Scott RJ, Butzkueven H, et al. Epigenome-wide association studies: current knowledge, strategies and recommendations. *Clin Epigenetics* [Internet]. 2021 Dec 1 [cited 2023 Sep 5];13(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/34863305/>
 88. García-Giménez JL, Seco-Cervera M, Tollefsbol TO, Romá-Mateo C, Peiró-Chova L, Lapunzina P, et al. Epigenetic biomarkers: Current strategies and future challenges for their use in the clinical laboratory. *Crit Rev Clin Lab Sci* [Internet]. 2017 Nov 17 [cited 2023 Oct 19];54(7–8):529–50. Available from: <https://pubmed.ncbi.nlm.nih.gov/29226748/>
 89. Maksimovic J, Phipson B, Oshlack A. A cross-package Bioconductor workflow for analysing methylation array data. *F1000Research* [Internet]. 2017 [cited 2020 Nov 17];5. Available from: <https://pubmed.ncbi.nlm.nih.gov/27347385/>
 90. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* [Internet]. 2013 Jan 15 [cited 2023 Aug 16];29(2):189–96. Available from: <https://pubmed.ncbi.nlm.nih.gov/23175756/>
 91. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol* [Internet]. 2012 [cited 2023 Aug 16];13(6). Available from: <https://pubmed.ncbi.nlm.nih.gov/22703947/>
 92. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. *Epigenomics* [Internet]. 2011 Dec [cited 2023 Aug 16];3(6):771–84. Available from: <https://pubmed.ncbi.nlm.nih.gov/22126295/>
 93. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol* [Internet]. 2014 Dec

- 3 [cited 2023 Aug 16];15(12). Available from: <https://pubmed.ncbi.nlm.nih.gov/25599564/>
94. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, et al. ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics* [Internet]. 2014 Feb 1 [cited 2023 Aug 16];30(3):428–30. Available from: <https://pubmed.ncbi.nlm.nih.gov/24336642/>
95. Price EM, Robinson WP. Adjusting for Batch Effects in DNA Methylation Microarray Data, a Lesson Learned. *Front Genet* [Internet]. 2018 Mar 16 [cited 2023 Aug 17];9(MAR). Available from: </pmc/articles/PMC5864890/>
96. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* [Internet]. 2010 Nov 30 [cited 2023 Nov 14];11(1):1–9. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-587>
97. Jaffe AE, Murakami P, Lee H, Leek JT, Fallin MD, Feinberg AP, et al. Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *Int J Epidemiol* [Internet]. 2012 Feb [cited 2023 Aug 18];41(1):200–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/22422453/>
98. Mallik S, Odom GJ, Gao Z, Gomez L, Chen X, Wang L. An evaluation of supervised methods for identifying differentially methylated regions in Illumina methylation arrays. *Brief Bioinform* [Internet]. 2019 Nov 1 [cited 2023 Aug 18];20(6):2224–35. Available from: <https://pubmed.ncbi.nlm.nih.gov/30239597/>
99. Pedersen BS, Schwartz DA, Yang I V., Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics* [Internet]. 2012 Nov [cited 2023 Aug 18];28(22):2986–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/22954632/>
100. Blake JA, Christie KR, Dolan ME, Drabkin HJ, Hill DP, Ni L, et al. Gene

- Ontology Consortium: going forward. *Nucleic Acids Res* [Internet]. 2015 Jan 28 [cited 2023 Aug 18];43(Database issue):D1049–56. Available from: <https://pubmed.ncbi.nlm.nih.gov/25428369/>
101. KEGG: Kyoto Encyclopedia of Genes and Genomes [Internet]. [cited 2021 Jan 1]. Available from: <https://www.genome.jp/kegg/>
 102. Kuleshov M V., Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* [Internet]. 2016 Jul 8 [cited 2023 Jan 17];44(W1):W90–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/27141961/>
 103. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. Gibas C, editor. *PLoS One* [Internet]. 2011 Jul 18 [cited 2021 Jan 8];6(7):e21800. Available from: <https://dx.plos.org/10.1371/journal.pone.0021800>
 104. Gene Ontology Resource [Internet]. [cited 2021 Jan 1]. Available from: <http://geneontology.org/>
 105. STRING: functional protein association networks [Internet]. [cited 2021 Jul 27]. Available from: https://string-db.org/cgi/about?sessionId=bUFZmne5zY0l&footer_active_subpage=references
 106. Otasek D, Morris JH, Bouças J, Pico AR, Demchak B. Cytoscape Automation: Empowering workflow-based network analysis. *Genome Biol*. 2019 Sep 2;20(1).
 107. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, et al. DAVID Bioinformatics Resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res*. 2007 Jul;35(SUPPL.2).
 108. Thompson M, Hill BL, Rakocz N, Chiang JN, Geschwind D, Sankararaman S, et al. Methylation risk scores are associated with a collection of phenotypes within electronic health record systems. *NPJ genomic Med* [Internet]. 2022 Dec 1 [cited 2023 Jul 6];7(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/36008412/>

109. Paudel D, Dhakal S, Parajuli S, Adhikari L, Peng Z, Qian Y, et al. Use of quantitative trait loci to develop stress tolerance in plants. *Plant Life under Chang Environ Responses Manag.* 2020 Jan 1;9:17–65.
110. Smith AK, Kilaru V, Kocak M, Almlı LM, Mercer KB, Ressler KJ, et al. Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. *BMC Genomics* [Internet]. 2014 Feb 21 [cited 2023 Sep 5];15(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/24555763/>
111. Min JL, Hemani G, Hannon E, Dekkers KF, Castillo-Fernandez J, Luijk R, et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat Genet.* 2021 Sep;53(9):1311–21.
112. Pan H, Holbrook JD, Karnani N, Kwoh CK. Gene, Environment and Methylation (GEM): a tool suite to efficiently navigate large scale epigenome wide association studies and integrate genotype and interaction between genotype and environment. *BMC Bioinformatics* [Internet]. 2016 Aug 2 [cited 2023 Sep 5];17(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/27480116/>
113. Bonora E, Formentini G, Calcaterra F, Lombardi S, Marini F, Zenari L, et al. HOMA-Estimated Insulin Resistance Is an Independent Predictor of Cardiovascular Disease in Type 2 Diabetic Subjects Prospective data from the Verona Diabetes Complications Study. *Diabetes Care* [Internet]. 2002 Jul 1 [cited 2023 Nov 13];25(7):1135–41. Available from: <https://dx.doi.org/10.2337/diacare.25.7.1135>
114. Child growth standards [Internet]. [cited 2023 Nov 14]. Available from: <https://www.who.int/tools/child-growth-standards>
115. Jenum AK, Mrøkrıd K, Sletner L, Vange S, Torper JL, Nakstad B, et al. Impact of ethnicity on gestational diabetes identified with the WHO and the modified International Association of Diabetes and Pregnancy Study Groups criteria: a population-based cohort study. *Eur J Endocrinol* [Internet]. 2012 Feb 1 [cited 2023 Jan 16];166(2):317–24. Available from: <https://eje.bioscientifica.com/view/journals/eje/166/2/317.xml>
116. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP,

- Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* [Internet]. 2014 May 15 [cited 2023 Aug 16];30(10):1363–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/24478339/>
117. Michels KB, Binder AM, Dedeurwaerder S, Epstein CB, Greally JM, Gut I, et al. Recommendations for the design and analysis of epigenome-wide association studies. *Nat Methods* [Internet]. 2013 Oct 27 [cited 2020 Apr 15];10(10):949–55. Available from: <http://www.nature.com/articles/nmeth.2632>
 118. Alptekin H, Çizmeciöäyü A, Işık H, Cengiz T, Yildiz M, Iyisoy MS. Predicting gestational diabetes mellitus during the first trimester using anthropometric measurements and HOMA-IR. *J Endocrinol Invest* [Internet]. 2016 May 1 [cited 2023 Mar 29];39(5):577–83. Available from: <https://pubmed.ncbi.nlm.nih.gov/26754418/>
 119. Fraszczyk E, Spijkerman AMW, Zhang Y, Brandmaier S, Day FR, Zhou L, et al. Epigenome-wide association study of incident type 2 diabetes: a meta-analysis of five prospective European cohorts. *Diabetologia* [Internet]. 2022 May 1 [cited 2022 Dec 2];65(5):763–76. Available from: <https://pubmed.ncbi.nlm.nih.gov/35169870/>
 120. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. *J Stat Softw* [Internet]. 2017 Dec 6 [cited 2022 Dec 2];82(13):1–26. Available from: <https://www.jstatsoft.org/index.php/jss/article/view/v082i13>
 121. Myers TA, Chanock SJ, Machiela MJ. LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations. *Front Genet* [Internet]. 2020 Feb 28 [cited 2023 Jan 17];11. Available from: <https://pubmed.ncbi.nlm.nih.gov/32180801/>
 122. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* [Internet]. 2019 Nov 1 [cited 2023 Apr 27];35(22):4851–3. Available from: <https://pubmed.ncbi.nlm.nih.gov/31233103/>

123. Moen GH, Sommer C, Prasad RB, Sletner L, Groop L, Qvigstad E, et al. Epigenetic modifications and gestational diabetes: A Systematic review of published literature [Internet]. Vol. 176, *European Journal of Endocrinology*. BioScientifica Ltd.; 2017 [cited 2020 Sep 28]. p. R247–67. Available from: www.eje-online.org
124. Wu P, Farrell WE, Haworth KE, Emes RD, Kitchen MO, Glossop JR, et al. Maternal genome-wide DNA methylation profiling in gestational diabetes shows distinctive disease-associated changes relative to matched healthy pregnancies. *Epigenetics*. 2016 Mar;0.
125. Dias S, Adam S, Rheeder P, Louw J, Pfeiffer C. Altered Genome-Wide DNA Methylation in Peripheral Blood of South African Women with Gestational Diabetes Mellitus. *Int J Mol Sci* [Internet]. 2019 Nov 20 [cited 2020 Apr 20];20(23):5828. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31757015>
126. Kang J, Lee CN, Li HY, Hsu KH, Lin SY. Genome-wide DNA methylation variation in maternal and cord blood of gestational diabetes population. *Diabetes Res Clin Pract*. 2017 Oct;132:127–36.
127. Ballesteros M, Gil-Lluís P, Ejarque M, Diaz-Perdigones C, Martinez-Guasch L, Fernández-Veledo S, et al. DNA methylation in gestational diabetes and its predictive value for postpartum glucose disturbances. *J Clin Endocrinol Metab* [Internet]. 2022 Aug 2 [cited 2022 Aug 9]; Available from: <https://pubmed.ncbi.nlm.nih.gov/35914803/>
128. S D, S A, P R, J L, C P. Altered Genome-Wide DNA Methylation in Peripheral Blood of South African Women with Gestational Diabetes Mellitus. *Int J Mol Sci* [Internet]. 2019 Dec 1 [cited 2021 Sep 22];20(23). Available from: <https://pubmed.ncbi.nlm.nih.gov/31757015/>
129. Doria A, Patti ME, Kahn CR. The Emerging Genetic Architecture of Type 2 Diabetes. *Cell Metab* [Internet]. 2008 Sep 3 [cited 2022 Aug 10];8(3):186–200. Available from: <http://www.cell.com/article/S1550413108002477/fulltext>
130. Florez JC. The Genetics of Type 2 Diabetes: A Realistic Appraisal in

2008. *J Clin Endocrinol Metab* [Internet]. 2008 Dec 1 [cited 2022 Aug 10];93(12):4633–42. Available from: <https://academic.oup.com/jcem/article/93/12/4633/2627272>
131. Staiger H, Machicao F, Fritsche A, Häring HU. Pathomechanisms of Type 2 Diabetes Genes. *Endocr Rev* [Internet]. 2009 Oct 1 [cited 2022 Aug 10];30(6):557–85. Available from: <https://academic.oup.com/edrv/article/30/6/557/2355053>
132. Mollet IG, Malm HA, Wendt A, Orho-Melander M, Eliasson L. Integrator of Stress Responses Calmodulin Binding Transcription Activator 1 (Camta1) Regulates miR-212/miR-132 Expression and Insulin Secretion. *J Biol Chem* [Internet]. 2016 Aug 26 [cited 2022 Aug 10];291(35):18440–52. Available from: <https://pubmed.ncbi.nlm.nih.gov/27402838/>
133. Saltiel AR, Kahn CR. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* [Internet]. 2001 Dec 13 [cited 2023 Mar 29];414(6865):799–806. Available from: <https://pubmed.ncbi.nlm.nih.gov/11742412/>
134. Wei FJ, Cai CY, Yu P, Lv J, Ling C, Shi WT, et al. Quantitative candidate gene association studies of metabolic traits in Han Chinese type 2 diabetes patients. *Genet Mol Res* [Internet]. 2015 Nov 30 [cited 2022 Aug 10];14(4):15471–81. Available from: <https://pubmed.ncbi.nlm.nih.gov/26634513/>
135. Mohamed SA, Fernandez-Tajes J, Franks PW, Bennet L. GWAS in people of Middle Eastern descent reveals a locus protective of kidney function—a cross-sectional study. *BMC Med* [Internet]. 2022 Dec 1 [cited 2022 Aug 10];20(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/35227251/>
136. Cauchi S, Proença C, Choquet H, Gaget S, De Graeve F, Marre M, et al. Analysis of novel risk loci for type 2 diabetes in a general French population: the D.E.S.I.R. study. *J Mol Med (Berl)* [Internet]. 2008 Mar [cited 2022 Aug 10];86(3):341–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/18210030/>
137. Wang HL, Wang L, Zhao CY, Lan HY. Role of TGF-Beta Signaling in Beta Cell Proliferation and Function in Diabetes. *Biomolecules*

[Internet]. 2022 Mar 1 [cited 2022 Sep 26];12(3):373. Available from: /pmc/articles/PMC8945211/

138. Kobayashi H, Looker HC, Satake E, D'Addio F, Wilson JM, Saulnier PJ, et al. Neuroblastoma suppressor of tumorigenicity 1 is a circulating protein associated with progression to end-stage kidney disease in diabetes. *Sci Transl Med* [Internet]. 2022 Aug 10 [cited 2022 Sep 26];14(657):eabj2109. Available from: <https://www.science.org/doi/10.1126/scitranslmed.abj2109>
139. Chen Y, Ma B, Wang X, Zha X, Sheng C, Yang P, et al. Potential Functions of the BMP Family in Bone, Obesity, and Glucose Metabolism. *J Diabetes Res* [Internet]. 2021 [cited 2023 Oct 25];2021. Available from: /pmc/articles/PMC8249130/
140. Pankow JS, Decker PA, Berardi C, Hanson NQ, Sale M, Tang W, et al. Circulating cellular adhesion molecules and risk of diabetes: the Multi-Ethnic Study of Atherosclerosis (MESA). *Diabet Med*. 2016 Jul;33(7):985.
141. Pletsch-Borba L, Watzinger C, Fortner RT, Katzke V, Schwingshackl L, Sowah SA, et al. Biomarkers of Vascular Injury and Type 2 Diabetes: A Prospective Study, Systematic Review and Meta-Analysis. *J Clin Med*. 2019;8(12).
142. Ruiz-Ojeda FJ, Olza J, Gil Á, Aguilera CM. Oxidative stress and inflammation in obesity and metabolic syndrome. *Obes Oxidative Stress Diet Antioxidants*. 2018 Jan;1–15.
143. Antoun E, Kitaba NT, Titcombe P, Dalrymple K V., Garratt ES, Barton SJ, et al. Maternal dysglycaemia, changes in the infant's epigenome modified with a diet and physical activity intervention in pregnancy: Secondary analysis of a randomised control trial. *PLoS Med*. 2020 Nov 5;17(11).
144. Mulder RH, Neumann A, Cecil CAM, Walton E, Houtepen LC, Simpkin AJ, et al. Epigenome-wide change and variation in DNA methylation in childhood: trajectories from birth to late adolescence. *Hum Mol Genet*. 2021 Jan;30(1):119.
145. Bohlin J, Håberg SE, Magnus P, Reese SE, Gjessing HK, Magnus MC,

- et al. Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biol.* 2016 Oct;17(1).
146. Sharp GC, Salas LA, Monnereau C, Allard C, Yousefi P, Everson TM, et al. Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation: Findings from the pregnancy and childhood epigenetics (PACE) consortium. *Hum Mol Genet* [Internet]. 2017 Oct 15 [cited 2020 Sep 21];26(20):4067–85. Available from: <https://pubmed.ncbi.nlm.nih.gov/29016858/>
 147. Shorey-Kendrick LE, McEvoy CT, O’Sullivan SM, Milner K, Vuylsteke B, Tepper RS, et al. Impact of vitamin C supplementation on placental DNA methylation changes related to maternal smoking: association with gene expression and respiratory outcomes. *Clin Epigenetics.* 2021 Dec;13(1).
 148. England-Mason G, Merrill SM, Moore S, Giesbrecht GF, Letourneau N, MacDonald A, et al. Prenatal exposure to phthalates and DNA methylation in infant blood and buccal samples: an epigenome-wide association study. *prep.* 2021 May;163:107183.
 149. Ustinova M, Peculis R, Rescenko R, Rovite V, Zaharenko L, Elbere I, et al. Novel susceptibility loci identified in a genome-wide association study of type 2 diabetes complications in population of Latvia. *BMC Med Genomics.* 2021 Dec;14(1).
 150. Balakrishnan A, Guruprasad KP, Satyamoorthy K, Joshi MB. Interleukin-6 determines protein stabilization of DNA methyltransferases and alters DNA promoter methylation of genes associated with insulin signaling and angiogenesis. *Lab Investig* [Internet]. 2018 Sep 28 [cited 2020 Apr 17];98(9):1143–58. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29955086>
 151. Liu Y, Geng H, Duan B, Yang X, Ma A, Ding X. Identification of Diagnostic CpG Signatures in Patients with Gestational Diabetes Mellitus via Epigenome-Wide Association Study Integrated with Machine Learning. *Biomed Res Int.* 2021;2021.
 152. Li E, Luo T, Wang Y. Identification of diagnostic biomarkers in patients with gestational diabetes mellitus based on transcriptome gene expression and methylation correlation analysis. *Reprod Biol*

Endocrinol. 2019 Dec;17(1).

153. Wang X, Huang J, Zheng Y, Long S, Lin H, Zhang N, et al. Study on the relationship between DNA methylation of target CpG sites in peripheral blood and gestational diabetes during early pregnancy. *Sci Rep*. 2021 Dec;11(1).
154. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* [Internet]. 2017 Nov 1 [cited 2023 Jan 17];66(11):2888–902. Available from: <https://pubmed.ncbi.nlm.nih.gov/28566273/>
155. Adeyemo A, Gerry N, Chen G, Herbert A, Doumatey A, Huang H, et al. A genome-wide association study of hypertension and blood pressure in African Americans. *PLoS Genet* [Internet]. 2009 Jul [cited 2023 Jan 17];5(7). Available from: <https://pubmed.ncbi.nlm.nih.gov/19609347/>
156. Liu JZ, Van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* [Internet]. 2015 Aug 27 [cited 2023 Jan 17];47(9):979–86. Available from: <https://pubmed.ncbi.nlm.nih.gov/26192919/>
157. Moccia C, Popovic M, Isaevska E, Fiano V, Trevisan M, Rusconi F, et al. Birthweight DNA methylation signatures in infant saliva. *Clin Epigenetics* [Internet]. 2021 Dec 1 [cited 2023 Sep 6];13(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/33741061/>
158. Sharma R, Frasn MG, Zelgert C, Zimmermann P, Fabre B, Wilson R, et al. Maternal-fetal stress and DNA methylation signatures in neonatal saliva: an epigenome-wide association study. *Clin Epigenetics* [Internet]. 2022 Dec 1 [cited 2023 Sep 6];14(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/35836289/>
159. FARP2 FERM, ARH/RhoGEF and pleckstrin domain protein 2 [Homo sapiens (human)] - Gene - NCBI [Internet]. [cited 2023 Sep 11]. Available from: <https://www.ncbi.nlm.nih.gov/gene/9855#summary>

160. Sinnott-Armstrong N, Tanigawa Y, Amar D, Mars N, Benner C, Aguirre M, et al. Genetics of 35 blood and urine biomarkers in the UK Biobank. *Nat Genet* [Internet]. 2021 Feb 1 [cited 2023 Sep 11];53(2):185–94. Available from: <https://pubmed.ncbi.nlm.nih.gov/33462484/>
161. Graham SE, Clarke SL, Wu KHH, Kanoni S, Zajac GJM, Ramdas S, et al. The power of genetic diversity in genome-wide association studies of lipids. *Nature* [Internet]. 2021 Dec 23 [cited 2023 Sep 12];600(7890):675–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/34887591/>
162. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshihara S, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet* [Internet]. 2021 Oct 1 [cited 2023 Sep 12];53(10):1415–24. Available from: <https://pubmed.ncbi.nlm.nih.gov/34594039/>
163. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* [Internet]. 2015 Feb 12 [cited 2020 Sep 8];518(7538):197–206. Available from: <https://www.nature.com/articles/nature14177>
164. Koskeridis F, Evangelou E, Said S, Boyle JJ, Elliott P, Dehghan A, et al. Pleiotropic genetic architecture and novel loci for C-reactive protein levels. *Nat Commun* [Internet]. 2022 Dec 1 [cited 2023 Sep 12];13(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/36376304/>
165. Yengo L, Vedantam S, Marouli E, Sidorenko J, Bartell E, Sakaue S, et al. A saturated map of common genetic variants associated with human height. *Nature* [Internet]. 2022 Oct 27 [cited 2023 Sep 12];610(7933):704–12. Available from: <https://pubmed.ncbi.nlm.nih.gov/36224396/>
166. BRINP2 BMP/retinoic acid inducible neural specific 2 [Homo sapiens (human)] - Gene - NCBI [Internet]. [cited 2023 Sep 12]. Available from: <https://www.ncbi.nlm.nih.gov/gene/57795#summary>
167. Wang H, Zhang F, Zeng J, Wu Y, Kemper KE, Xue A, et al. Genotype-

- by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. *Sci Adv* [Internet]. 2019 Aug 14 [cited 2023 Sep 12];5(8). Available from: <https://pubmed.ncbi.nlm.nih.gov/31453325/>
168. Lind L. Genome-Wide Association Study of the Metabolic Syndrome in UK Biobank. *Metab Syndr Relat Disord* [Internet]. 2019 Dec 1 [cited 2023 Sep 12];17(10):505–11. Available from: <https://pubmed.ncbi.nlm.nih.gov/31589552/>
169. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* [Internet]. 2019 Jun 27 [cited 2023 Sep 12];570(7762):514–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/31217584/>
170. Huang J, Huffman JE, Huang Y, Do Valle Í, Assimes TL, Raghavan S, et al. Genomics and phenomics of body mass index reveals a complex disease network. *Nat Commun* [Internet]. 2022 Dec 1 [cited 2023 Sep 12];13(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/36581621/>
171. Hollis B, Day FR, Busch AS, Thompson DJ, Soares ALG, Timmers PRHJ, et al. Genomic analysis of male puberty timing highlights shared genetic basis with hair colour and lifespan. *Nat Commun* [Internet]. 2020 Dec 1 [cited 2023 Sep 12];11(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/32210231/>
172. Riveros-McKay F, Mistry V, Bounds R, Hendricks A, Keogh JM, Thomas H, et al. Genetic architecture of human thinness compared to severe obesity. *PLoS Genet* [Internet]. 2019 [cited 2023 Sep 12];15(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/30677029/>
173. Wu P, Farrell WE, Haworth KE, Emes RD, Kitchen MO, Glossop JR, et al. Maternal genome-wide DNA methylation profiling in gestational diabetes shows distinctive disease-associated changes relative to matched healthy pregnancies. *Epigenetics* [Internet]. 2018 Feb 1 [cited 2021 Oct 8];13(2):122. Available from: [/pmc/articles/PMC5873366/](https://pubmed.ncbi.nlm.nih.gov/30677029/)

174. Opsahl JO, Moen GH, Qvigstad E, Böttcher Y, Birkeland KI, Sommer C. Epigenetic signatures associated with maternal body mass index or gestational weight gain: a systematic review. *J Dev Orig Health Dis* [Internet]. 2021 Jun 1 [cited 2023 May 15];12(3):373–83. Available from: <https://pubmed.ncbi.nlm.nih.gov/32873364/>

175. Cardona A, Day FR, Perry JRB, Loh M, Chu AY, Lehne B, et al. Epigenome-Wide Association Study of Incident Type 2 Diabetes in a British Population: EPIC-Norfolk Study. *Diabetes* [Internet]. 2019 Dec 1 [cited 2023 Feb 15];68(12):2315–26. Available from: <https://diabetesjournals.org/diabetes/article/68/12/2315/39835/Epigenome-Wide-Association-Study-of-Incident-Type>

176. Meeks KAC, Henneman P, Venema A, Addo J, Bahendeka S, Burr T, et al. Epigenome-wide association study in whole blood on type 2 diabetes among sub-Saharan African individuals: findings from the RODAM study. *Int J Epidemiol* [Internet]. 2019 Feb 1 [cited 2023 Feb 3];48(1):58–70. Available from: <https://pubmed.ncbi.nlm.nih.gov/30107520/>

177. Tobi EW, Juvinao-Quintero DL, Ronkainen J, Ott R, Alfano R, Canouil M, et al. Maternal Glycemic Dysregulation During Pregnancy and Neonatal Blood DNA Methylation: Meta-analyses of Epigenome-Wide Association Studies. *Diabetes Care* [Internet]. 2022 Mar 1 [cited 2023 Feb 9];45(3):614–23. Available from: <https://pubmed.ncbi.nlm.nih.gov/35104326/>

178. Qie R, Chen Q, Wang T, Chen X, Wang J, Cheng R, et al. Association of ABCG1 gene methylation and its dynamic change status with incident type 2 diabetes mellitus: the Rural Chinese Cohort Study. *J Hum Genet* [Internet]. 2021 Apr 1 [cited 2023 Feb 9];66(4):347–57. Available from: <https://pubmed.ncbi.nlm.nih.gov/32968204/>

179. Tian M, Zhang C, Han H, Liu R, Gao Z, Chen P, et al. Effects of the preassembly of benzohydroxamic acid with Fe (III) ions on its adsorption on cassiterite surface. *Miner Eng* [Internet]. 2018 Oct 1 [cited 2018 Nov 13];127:32–41. Available from: <https://www.sciencedirect.com/science/article/pii/S0892687518303352>

180. MAP4K2 mitogen-activated protein kinase kinase kinase kinase 2

[Homo sapiens (human)] - Gene - NCBI [Internet]. [cited 2023 Feb 9]. Available from: <https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=5871>

181. PDGFB platelet derived growth factor subunit B [Homo sapiens (human)] - Gene - NCBI [Internet]. [cited 2023 Feb 13]. Available from: <https://www.ncbi.nlm.nih.gov/gene/5155#gene-ontology>
182. Shan Z, Xu C, Wang W, Li W. Enhanced PDGF signaling in gestational diabetes mellitus is involved in pancreatic β -cell dysfunction. *Biochem Biophys Res Commun*. 2019 Aug 20;516(2):402–7.
183. CG H, B C, R F, J J, T K, S L, et al. Maternal Gestational Diabetes Mellitus and Newborn DNA Methylation: Findings From the Pregnancy and Childhood Epigenetics Consortium. *Diabetes Care* [Internet]. 2020 Jan 1 [cited 2021 Sep 22];43(1):98–105. Available from: <https://pubmed.ncbi.nlm.nih.gov/31601636/>
184. Canouil M, Khamis A, Keikkala E, Hummel S, Lobbens S, Bonnefond A, et al. Epigenome-Wide Association Study Reveals Methylation Loci Associated With Offspring Gestational Diabetes Mellitus Exposure and Maternal Methylome. *Diabetes Care* [Internet]. 2021 Sep 1 [cited 2023 Mar 29];44(9):1992–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/34116986/>

Material suplementario

Tablas suplementarias:

Tabla Suplementaria 1- 42 CpGs asociados a la DM2 en el meta-análisis, utilizados para crear el MRS.

llumina ID	Gene name	CHR	Position	Effect size	Standard Error	HetISq	HetPVal	P.value	FDR	Direction across studies ^a	Effect size replication	P.value replication
cg19693031	<i>TXNIP</i>	1	145441552	-0.0199	0.002	87,8	1,347E-	3.89E-23	1.84E-17	-----	-0.0119	1.15E-12
cg06500161	<i>ABCG1</i>	21	43656587	0.0109	0.0011	63,2	0,028	9.26E-22	2.19E-16	+++++	0.0092	7.58E-20
cg11024682	<i>SREBF1</i>	17	17730094	0.0088	0.0011	42,7	0,137	1.41E-14	2.23E-09	+++++	0.0075	1.28E-13
cg00574958	<i>CPT1A</i>	11	68607622	-0.0051	0.0007	80,5	0,000	6.35E-13	7.52E-08	-----	-0.0037	1.99E-07
cg05778424	<i>AKAP1</i>	17	55169508	0.0075	0.0011	65	0,022	3.56E-11	3.37E-06	+++++	0.0043	7.07E-06
cg14020176	<i>SLC9A3R1</i>	17	72764985	0.0087	0.0014	33,9	0,209	1.53E-10	1.21E-05	+?+++	0.0033	3.63E-04
cg07504977	<i>OLMALINC</i>	10	102131012	0.0112	0.0018	0	0,716	2.17E-10	1.47E-05	+++++	0.0062	1.37E-04
cg04816311	<i>C7orf50</i>	7	1066650	0.0112	0.0018	66,5	0,018	2.76E-10	1.60E-05	+++++	0.0059	1.72E-03
cg14476101	<i>PHGDH</i>	1	120255992	-0.0145	0.0023	59,9	0,041	3.04E-10	1.60E-05	-----	-0.01	2.15E-06
cg27243685	<i>ABCG1</i>	21	43642366	0.0061	0.001	0	0,801	4.26E-10	2.02E-05	+++++	0.0052	8.45E-13
cg14870271	<i>LGALS3BP</i>	17	76976010	0.0084	0.0014	39	0,161	5.00E-10	2.15E-05	+++++	0.0056	4.98E-05
cg26262157	<i>PFKFB3</i>	10	6214079	-0.009	0.0015	63,9	0,026	1.03E-09	3.47E-05	-----	-0.0054	1.05E-04
cg18568872	<i>ZNF710</i>	15	90606494	0.006	0.001	0	0,587	1.22E-09	3.84E-05	+++++	0.0029	3.69E-04
cg01373896	<i>KLF16</i>	19	1854724	0.0068	0.0011	50,1	0,091	1.30E-09	3.84E-05	+++++	0.002	5.45E-02
cg08994060	<i>PFKFB3</i>	10	6214026	-0.0107	0.0018	62	0,032	1.66E-09	4.46E-05	-----	-0.0056	1.11E-03
cg11202345	<i>LGALS3BP</i>	17	76976057	0.0078	0.0013	0	0,496	1.70E-09	4.46E-05	+++++	0.0062	6.55E-06
cg21480264	<i>POLN</i>	4	2137264	0.0059	0.001	0	0,826	2.45E-09	6.12E-05	+++++	0.0025	2.73E-03

cg19750657	<i>UFM1</i>	13	38935967	0.0089	0.0015	15,7	0,314	2.66E-09	6.16E-05	+++++	0.0065	7.09E-07
cg06192883	<i>MYO5C</i>	15	52554171	0.008	0.0013	80,5	3,95E-	2.73E-09	6.16E-05	+++++	0.0066	4.53E-06
cg06378491	<i>MAP4K2</i>	11	64564012	0.0044	0.0008	47,1	0,109	3.45E-09	7.11E-05	+++++	0.0025	9.22E-05
cg11269166	<i>METTL8</i>	2	172203847	0.0066	0.0011	80,9	3,28E-	4.19E-09	7.94E-05	+++++	0.0025	2.31E-02
cg08788930	<i>DENND3</i>	8	142201685	0.0073	0.0013	37,7	0,170	5.83E-09	0.000106	+++++	0.0034	1.52E-03
cg22650271	<i>SYNGR1</i>	22	39760165	0.0057	0.001	72,4	0,006	6.62E-09	0.000116	+++++	0.0034	4.62E-04
cg03691549	<i>TENC1</i>	12	53443911	0.0059	0.001	0	0,674	9.51E-09	0.000161	+++++	0.0026	8.53E-03
cg09664445	<i>KIAA0664</i>	17	2612406	0.0054	0.0009	37,6	0,171	1.03E-08	0.000163	+++++	0.0021	1.48E-02
cg14956201	<i>TRIO</i>	5	14358153	0.0077	0.0014	0	0,493	1.11E-08	0.00017	+++++	0.0013	3.24E-01
cg17540192	<i>TECPR1</i>	7	97875259	0.0048	0.0008	71,4	0,007	1.47E-08	0.000218	+++++	0.0027	1.28E-03
cg25217710	<i>BCAN</i>	1	156609523	0.0052	0.0009	67,6	0,015	1.52E-08	0.000218	+++++	0.0027	2.97E-04
cg16861241	<i>FOXJ1</i>	17	74138396	0.0056	0.001	0	0,782	1.65E-08	0.00023	+++++	0.0005	5.84E-01
cg15020801	<i>PNPO</i>	17	46022809	0.007	0.0012	40,9	0,149	1.80E-08	0.000244	+++++	0.006	4.95E-07
cg02879453	<i>ADCY7</i>	16	50321818	0.0078	0.0014	50,8	0,087	3.99E-08	0.000525	+++++	0.0054	1.17E-05
cg04682775	<i>SLC6A9</i>	1	44495089	0.0066	0.0012	8,3	0,359	4.17E-08	0.000534	+++++	0.0011	2.55E-01
cg10639435	<i>ZNF250</i>	8	146104221	0.0077	0.0014	0	0,745	4.28E-08	0.000534	+++++	0.0014	3.06E-01
cg04927537	<i>LGALS3BP</i>	17	76976091	0.0103	0.0019	0	0,587	4.87E-08	0.000592	+++++	0.0063	4.68E-04
cg16097041	<i>FLAD1</i>	1	154965544	0.0061	0.0011	0	0,509	5.05E-08	0.000599	+++++	0.0031	7.41E-04
cg25130381	<i>SLC9A1</i>	1	27440721	0.0057	0.001	55,7	0,060	5.81E-08	0.000671	+++++	0.0049	6.85E-07
cg20507228	<i>MAN2A2</i>	15	91460071	0.0124	0.0023	0	0,809	6.37E-08	0.000719	+?+++	0.0048	2.12E-03
cg05460226	<i>PIK3R5</i>	17	8804279	0.0095	0.0018	15,2	0,318	6.92E-08	0.000763	+++++	NA	NA

cg01101459	<i>LINC01132</i>	1	234871477	0.0065	0.0012	74,5	0,003	7.45E-08	0.000802	+++++	0.0086	1.59E-09
cg25178683	<i>LGALS3BP</i>	17	76976267	0.0083	0.0015	0	0,471	8.06E-08	0.000848	+++++	0.0059	4.22E-05
cg09072148	<i>NRXN2</i>	11	64491639	0.0037	0.0007	24,1	0,261	9.11E-08	0.000939	+++++	0.0004	4.50E-01
cg12322877	<i>ASPSCR1</i>	17	79963213	0.0119	0.0022	71,4	0,007	9.70E-08	0.000978	-++++	0.0062	2.97E-03

Tabla Suplementaria 23- 20 CpGs más hipometilados y más hipermetilados en el array de madres en la visita inicial.

CpG_ID	logFC	P.Value	adj.P.Val	C_T0_AVG	DG_T0_AVG	CHR	MAPINFO	gene	feature	cgi
cg13824270	-0.307	1.10E-04	0.042	0.595	0.422	6	4020946	PRPF4B	TSS1500	shore
cg20409752	-0.264	1.49E-04	0.048	0.679	0.496	1	7122726	CAMTA1	Body	island
cg05091658	-0.253	2.67E-05	0.023	0.400	0.272	10	6954856		IGR	opensea
cg05373962	-0.247	7.81E-05	0.036	0.612	0.443	22	49881684		IGR	shore
cg03411475	-0.229	9.49E-06	0.015	0.429	0.339	10	29818369	SVIL	Body	opensea
cg26736055	-0.223	1.19E-05	0.017	0.866	0.757	14	100627507	DEGS2	TSS1500	shore
cg04736673	-0.216	1.56E-04	0.048	0.369	0.265	10	61647141	CCDC6	Body	opensea
cg18217622	-0.207	8.73E-10	0.000	0.563	0.420	6	170732253		IGR	island
cg15953550	-0.206	5.92E-05	0.032	0.747	0.616	2	217384279		IGR	opensea
cg01824150	-0.206	7.00E-05	0.034	0.555	0.424	17	47342475		IGR	opensea
cg02148796	-0.197	4.11E-06	0.010	0.653	0.544	2	29548577	ALK	Body	opensea
cg24158931	-0.191	6.26E-07	0.005	0.637	0.520	14	92116289	CATSPERB	Body	opensea
cg07257824	-0.181	5.62E-08	0.002	0.545	0.414	6	170732353		IGR	island
cg06134980	-0.177	6.89E-05	0.034	0.687	0.577	15	50551582	HDC	Body	opensea
cg18727120	-0.171	1.29E-04	0.045	0.789	0.648	17	406137		IGR	shore
cg12718519	-0.166	4.01E-05	0.027	0.376	0.247	1	2058417	PRKCZ	5'UTR	opensea
cg06733558	-0.166	1.98E-07	0.003	0.814	0.744	20	21216097	KIZ	Body	opensea
cg24529535	-0.163	1.77E-06	0.007	0.525	0.422	1	152321040	FLG-AS1	Body	opensea
cg00596184	-0.163	1.92E-05	0.021	0.778	0.707	16	69350052	VPS4A	Body	opensea
cg24083051	-0.156	9.11E-06	0.015	0.292	0.226	22	28181085	MN1	Body	opensea

cg10930308	0.161	4.22E-05	0.028	0.176	0.255	6	30039476	RNF39	Body	island
cg05433805	0.161	3.14E-06	0.009	0.306	0.422	3	37204814	LRRFIP2	5'UTR	opensea
cg09023351	0.161	1.34E-08	0.001	0.310	0.447	5	108791617		IGR	opensea
cg04217177	0.163	2.47E-05	0.023	0.514	0.574	1	217306700	ESRRG	5'UTR	shore
cg20891558	0.171	1.56E-04	0.048	0.234	0.323	2	74357851		IGR	island
cg26861333	0.172	2.31E-05	0.022	0.608	0.753	16	52631882	CASC16	Body	opensea
cg08256922	0.172	5.35E-06	0.012	0.635	0.708	6	31050292		IGR	opensea
cg20744362	0.175	5.50E-06	0.012	0.642	0.728	22	50050164	C22orf34	Body	opensea
cg13918804	0.180	1.13E-06	0.006	0.443	0.543	1	2043761	PRKCZ	5'UTR	shelf
cg01459453	0.182	1.06E-11	0.000	0.609	0.736	1	169599212	SELP	Body	opensea
cg25267487	0.210	4.24E-05	0.028	0.618	0.697	19	29217858		IGR	shore
cg03749207	0.212	1.44E-07	0.002	0.094	0.197	16	857863	PRR25	Body	island
cg14854723	0.217	6.27E-05	0.033	0.391	0.553	9	129202700	MVB12B	Body	opensea
cg23758309	0.228	3.11E-07	0.003	0.461	0.660	12	97273300		IGR	opensea
cg02938172	0.233	1.07E-05	0.016	0.458	0.566	17	185152	RPH3AL	5'UTR	island
cg14209346	0.234	9.31E-06	0.015	0.499	0.579	3	5625835		IGR	opensea
cg07253311	0.234	2.55E-05	0.023	0.651	0.788	2	175976580	ATF2	Body	opensea
cg24587835	0.234	1.42E-04	0.047	0.341	0.450	17	5674234	LOC339166	TSS1500	opensea
cg02518222	0.240	3.35E-05	0.025	0.614	0.762	4	88274874	HSD17B11	Body	opensea
cg20758756	0.274	1.51E-05	0.019	0.411	0.534	18	9321560		IGR	opensea
cg04248279	0.282	3.22E-05	0.025	0.581	0.722	17	184833	RPH3AL	5'UTR	shore

Tabla suplementaria 3- Resumen del número de DMPs según comparación establecida

Comparison	CpG sites adjusted by age, gestational age, pre-pregnant BMI, foetal sex, weigh gain, HOMA-IR and treatment (insulin or diet) *
GDM vs non-GDM at T0	757/384
GDM vs non-GDM at T1	311/154
T1 vs T0 in non-GDM	263/4
T1 vs T0 in GDM	0/0

*False Discover Rate (FDR) p-value <0.05. †Hypermethylated and Hypomethylated CpG sites

Tabla suplementaria 4- Validación de los CpGs en la cohorte EIPREG

EIPREG complete					EIPREG_EU					EIPREG_SA				
cg_ID	logFC	Media non-DG	Media DG	PValue	cg_ID	logFC	Media non-DG	Media DG	PValue	cg_ID	logFC	Media non-DG	Media DG	PValue
cg04095097	0.169	0.335	0.361	0.0042	cg21809624	-0.012	0.543	0.554	0.058	cg15589641	-0.025	0.473	0.463	0.024
cg27603605	0.071	0.669	0.681	0.016	cg04095097	0.0049	0.328	0.336	0.061	cg04095097	0.038	0.354	0.397	0.026
cg12080079	0.095	0.728	0.743	0.047	cg08386137	0.0049	0.852	0.858	0.243	cg18923740	-0.020	0.308	0.295	0.040
cg10102108	0.062	0.729	0.737	0.068	cg12080079	0.0096	0.727	0.736	0.250	cg14579430	-0.021	0.462	0.452	0.075
cg21809624	0.051	0.543	0.553	0.078	cg01757548	0.0051	0.908	0.909	0.270	cg04600077	0.0093	0.714	0.708	0.206
cg01743873	0.049	0.416	0.425	0.121	cg10102108	0.0054	0.725	0.73	0.374	cg01757548	0.0058	0.912	0.909	0.224
cg18923740	-0.067	0.297	0.289	0.171	cg14579430	0.008	0.455	0.444	0.456	cg10102108	0.0078	0.738	0.747	0.249
cg15589641	-0.051	0.466	0.458	0.240	cg12432693	-0.0019	0.898	0.821	0.487	cg07257824	-0.022	0.498	0.512	0.387
cg14579430	-0.054	0.457	0.447	0.256	cg04600077	0.0042	0.71	0.712	0.492	cg01743873	0.0059	0.445	0.443	0.422
cg23743013	-0.035	0.362	0.355	0.283	cg01743873	0.0038	0.404	0.413	0.557	cg14688342	-0.00519	0.674	0.74	0.442
cg07257824	-0.073	0.516	0.513	0.430	cg07257824	-0.0073	0.523	0.513	0.674	cg04802986	-0.0040	0.842	0.836	0.442
cg04802986	-0.024	0.839	0.837	0.493	cg18923740	0.0034	0.293	0.285	0.704	cg15329406	0.0058	0.302	0.294	0.445
cg04600077	-0.018	0.711	0.71	0.573	cg27603605	-0.0019	0.67	0.676	0.717	cg18217622	-0.015	0.49	0.513	0.473
cg14688342	-0.015	0.67	0.67	0.589	cg23743013	-0.0021	0.366	0.357	0.746	cg01459453	0.0082	0.751	0.761	0.504

cg08386137	0.018	0.854	0.856	0.593	cg18217622	-0.0049	0.514	0.506	0.748	cg06279296	0.014	0.261	0.278	0.514
cg06279296	-0.049	0.253	0.252	0.633	cg14688342	-0.0014	0.668	0.667	0.804	cg12215871	0.0056	0.811	0.81	0.545
cg18217622	-0.025	0.506	0.508	0.754	cg15329406	0.0015	0.286	0.285	0.814	cg21809624	0.0023	0.543	0.551	0.762
cg15329406	-0.0089	0.29	0.288	0.800	cg06279296	0.0042	0.25	0.234	0.824	cg12432693	-0.00099	0.9	0.899	0.768
cg01459453	-0.014	0.751	0.752	0.810	cg01459453	0.0015	0.752	0.746	0.873	cg27603605	0.0019	0.667	0.688	0.801
cg12215871	0.011	0.817	0.816	0.819	cg04802986	-220.00	0.838	0.836	0.885	cg12080079	-0.00087	0.728	0.753	0.928
cg01757548	0.0079	0.909	0.909	0.863	cg12215871	0.00082	0.82	0.821	0.899	cg23743013	-0.00073	0.354	0.355	0.931
cg12432693	0.00026	0.899	0.898	0.993	cg15589641	0.00077	0.463	0.456	0.936	cg08386137	-0.00025	0.858	0.853	0.960

logFC: log2 Fold Change. EPIPREG complete: análisis en toda la cohorte. EPIPREG_EU: análisis en la cohorte Europea. EPIPREG_SA: análisis en la Sudasiáticas.

Tabla suplementaria 5- CpGs y genes previamente relacionados con obesidad infantil utilizados para realizar el análisis.

Genes	FTO	HOXB5	KC6	KCTD15	MAST4	MTOR	NCKAP5L	NCOA1	PARD3	POMC	SEC16B	TMEM18
CpG_ID	cg18080211	cg00632306	cg16447539	cg05620842	cg11941228	cg13997649	cg11553795	cg07493520	cg03398910	cg06846259	cg05775049	cg26655295
	cg02168806	cg04196862	cg03641192	cg20952031	cg18175036	cg13435101	cg20418191	cg00774602	cg02801550	cg24718866	cg04408858	cg17084053
	cg04379260			cg14266436	cg02625902	cg21708354	cg16888824	cg25665767	cg14588422	cg20387815	cg06557390	cg05664523
	cg08855319			cg14608758	cg26172694	cg04751207	cg11605222	cg12407703	cg06841052	cg09527270	cg07088064	cg13955747
	cg06833284			cg26156860	cg18118497	cg11561443	cg22032626	cg15841062	cg15292765	cg00674304	cg25360898	cg07502926
	cg22885448			cg20638467	cg03883572	cg09004581	cg22047262	cg09669874	cg25322645	cg13025668	cg12256456	cg26378753
	cg26202110			cg00998135	cg16235531	cg15060694	cg01316762	cg10027944	cg06338697	cg10045137	cg25135333	cg10356655
	cg15817114			cg10721927	cg16630369	cg14586630	cg20108243	cg26165826	cg011176720	cg02757179	cg11455423	cg17611880
	cg12891543			cg11169405	cg11017382	cg14731191	cg12143106	cg10628062	cg00906319	cg24425171	cg02289040	cg08606075
	cg21724915			cg00704310	cg09648315	cg24990010	cg20446255	cg16469582	cg19568997	cg14357535	cg17383143	cg22046187
	cg26580413			cg07279094	cg20964285	cg15685954	cg00457650	cg09625143	cg07576664	cg02716646	cg04632468	cg16966133
	cg13947504			cg19189726	cg14546471	cg26201579	cg01000282	cg12246203	cg15828842	cg22900229	cg19695615	cg04521543
	cg09674325			cg08420232	cg09005125	cg14169841	cg08968229	cg10398589	cg04065742	cg23809645	cg24598929	cg06218177
	cg01207767			cg04566826	cg16230553	cg10549648	cg08354685	cg22372573	cg01113896	cg16302441	cg00482464	cg05060085
	cg21789594			cg02973307	cg02574484	cg09169367	cg13847622	cg03984919	cg07452799	cg09672383	cg02772103	cg24201303
cg02252501			cg16276290	cg12145624	cg19339848	cg14405137	cg18128666	cg05901114	cg02079741	cg06350805	cg00245746	

cg00714672			cg04542091	ch.5.1258129F	cg07977185	cg22126976	cg16584577	cg25212025	cg03560973	cg14937293	cg06200714
cg04036070			cg26149738	cg03539041	cg18357393	cg17212980	cg14849556	cg00859888	cg08030082	cg24688413	cg27237671
cg26357595			cg24707640	cg01933079	cg02388568	cg00874929	cg14337036	cg17392230	cg09916783	cg27451587	cg10560669
cg10499632			cg14919591	cg20691612	cg09414752	cg16145092	cg06022832	cg06546571	cg19241807	cg18039610	cg22071951
cg23634200			cg13613853	cg24689196	cg21765818	cg14438062	cg00609566	cg20120614	cg20807790	cg17971251	cg18263335
cg07349222			cg08143160	cg24947286	cg18727558	cg23171072	cg23756763	cg24138862	cg11894631	cg00647232	cg14918214
cg02576188			cg09786757	cg08743751	cg01061843	cg07805817	cg18546155	cg19884427	cg14170547	cg24974611	cg12368392
cg26982104			cg22625440	cg11831286	cg13636895	cg08606453	cg17733123	cg26662969	cg15715889	cg06306117	
cg02479755				cg06045225	cg17626203	cg25756241	cg23461314	cg25217535	cg01926269	cg08060884	
cg10227678				cg07510828	cg16179922	cg18818866	cg17034014	cg23804261	cg06904565		
cg02336817				cg11590213	cg06632353	cg04578169	cg18149011	cg09461005	cg00293936		
cg14051148				cg25557227	cg16128122	cg26038044	cg16114681	cg25405138			
cg13884987				cg08232855	cg21474053		cg20055284	cg06832449			
cg15762175				cg16214267	cg17255672		cg23648923	cg02538783			
cg09243909				cg00277936	cg03881436		cg08252585	cg23093462			
cg08551495				cg01056678	cg16708079			cg19192223			
cg16226789				cg17278401	cg08734093			cg07270259			
cg20983183				cg18919915	cg12146012			cg18351448			

cg13627320				cg02632185	cg19625377			cg15645287			
cg17412805				cg15386865	cg04310074			cg05709402			
cg20160665				cg25399309	cg11750962			cg27053542			
cg14641193				cg10665379	cg11141763			cg15726896			
cg02265089				cg19458529	cg22718696			cg16895672			
cg01193320				cg05961632	cg06831279			cg04421618			
cg15536925				cg09012346	cg26230245			cg06271128			
cg13927026				cg06003834	cg14535017			cg26988617			
cg09123899				cg04480129	cg12776698			cg27500214			
cg01130861				cg19751560	cg16520348			cg25874150			
cg19439737				ch.5.1264897R	cg00997327			cg24691332			
cg24136440				cg17449851	cg02717678			cg06564860			
cg09749256				cg27116842	cg08433465			cg19470652			
cg14994009				cg11809476	cg07493824			cg04685387			
cg21724020				cg02657823	cg26060021			cg24901662			
cg03318340				cg26967579	cg05403879			cg17200577			
cg26999702				cg01791798	cg01380153			cg06723644			
cg05888799				cg07956968	cg00104343			cg11161396			

cg00844553				cg05478243	cg15660755			cg26795540			
cg12016828				cg05134928	cg16118379			cg11830436			
cg06940838				cg13286857	cg06031849			cg09884107			
cg26695784				cg07416291	cg00909226			cg23548374			
cg10233583				cg15362075	cg07922924			cg09513647			
cg19284725				cg09672871	cg08248705			cg04903360			
cg02992067				cg14384984	cg06632603			cg13976324			
cg02111433				cg17462107	cg12750884			cg11646078			
cg19115653				cg09728102	cg19836803			cg10447404			
cg05305355				cg02107842	cg22968728			cg06479235			
cg26651810				cg12784701	cg23984342			cg01313176			
cg22348368				cg10629165	cg03784329			cg19017553			
cg03310510				cg19355078	cg16496699			cg14740976			
cg20324219				cg24155984	cg12100376			cg03310691			
cg00625110				cg24609937	cg07029998			cg09095047			
cg09789229				cg10810657	cg04343233			cg05877768			
cg07306006				cg11140679	cg21263609			cg13628006			
cg20561643				cg18786944	cg10222608			cg15523609			

cg13270001				cg02358190	cg14854718			cg20598295			
cg09861673				cg08727187	cg08302116			cg25017900			
cg01729592				cg03511806	cg22730007			cg19569531			
cg06434738				cg09872250	cg17291957			cg23220915			
cg00916441				cg22385596	cg15725046						
cg22311244				cg07034329	cg08862778						
cg02642561				cg13811092	cg20240601						
cg10331105				cg01243014	cg19821366						
cg03073269				cg14742628	cg24863624						
cg07562721				cg17606548	cg20001572						
cg00561634				cg06206313	cg09341685						
cg01485549				cg16113371	cg05677959						
cg04544898				cg05088356	cg09418268						
cg02173208				cg12568520	cg10315903						
cg27401543				cg12315942	cg13607512						
cg04711291				cg03947737							
cg14663066				cg01985733							
cg15164711				cg20396510							

cg03022104			cg09695117							
cg07873467			cg17760391							
cg01986630			cg08840173							
cg03312170			cg26460816							
cg03778087			cg13211731							
cg16649357			cg13695535							
cg05967340			cg10322906							
cg14868135			cg00480850							
cg06392242			cg19217739							
cg05587478			cg16642505							
cg00397287			cg10124928							
cg11297458			cg05036909							
			cg07546427							
			cg21122007							
			cg00164406							
			cg19544480							
			cg04375253							
			cg03254529							

				cg09719813							
				cg08023693							
				cg18745933							
				cg01297958							
				cg04599190							
				cg11567466							
				cg21158255							
				cg09568500							
				cg09525764							
				cg22724353							
				cg25373068							
				cg17686672							
				cg24811598							
				cg08274405							
				cg20986626							
				cg06123376							
				cg03872157							
				cg24464915							

				cg17578969							
				cg14455887							
				cg08244784							
				cg00068571							
				cg22565782							
				cg15808348							
				cg14105271							
				cg06468344							
				cg14583126							
				cg26444429							
				cg03005143							
				cg10012517							
				cg06040990							
				cg12264517							
				cg17723542							
				cg27596513							

Figuras suplementarias:

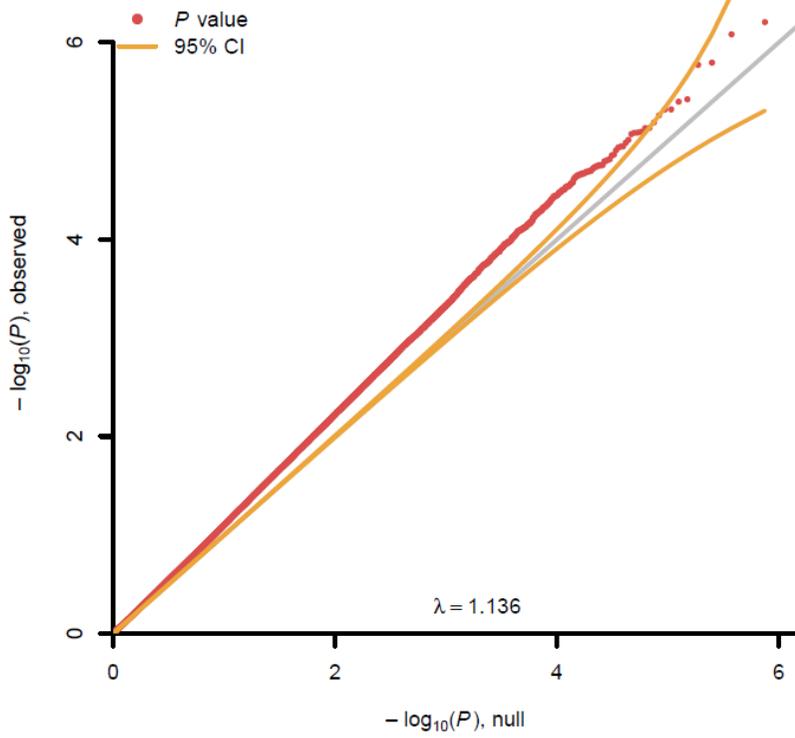


Figura suplementaria 3- QQplot del EWAS de sangre materna en la visita inicial. λ : Representa el valor de la inflación de los P value obtenidos, a mayor inflación mayor probabilidad de encontrar falsos positivos. $\lambda = 1.136$ inflación moderada.

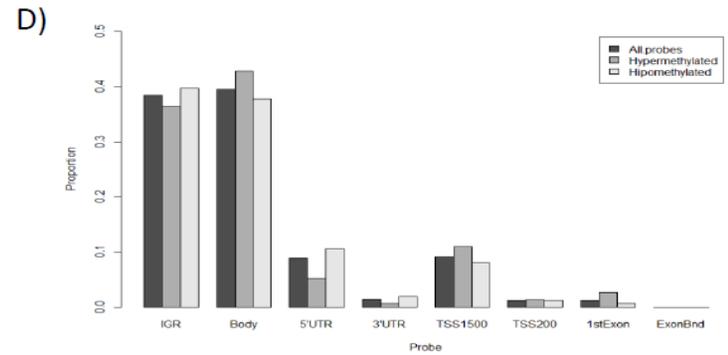
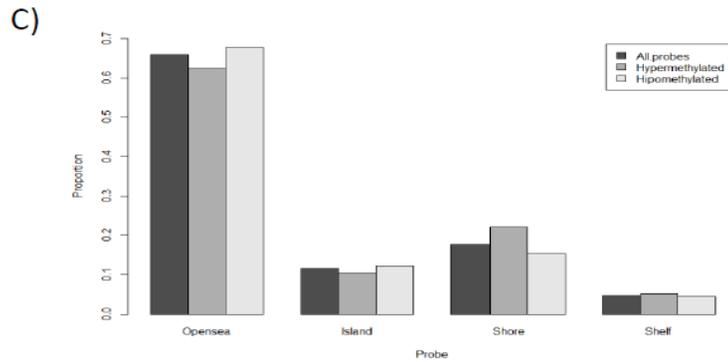
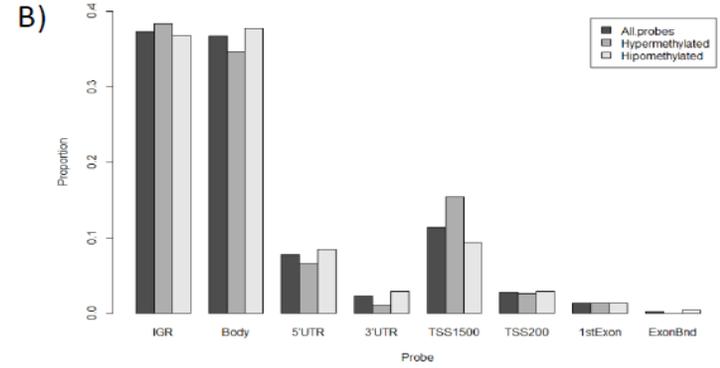
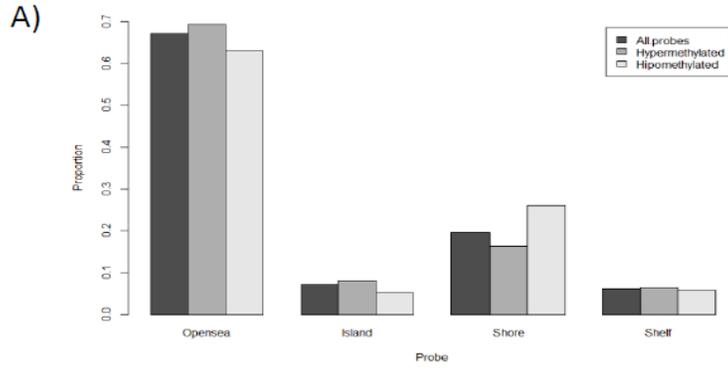


Figura suplementaria 2- Distribución de los CpGs en la visita basal (A, B) en la visita preparto (C, D), basado en la posición respecto a las islas CpGs (A, C); basado en la posición con respecto al gen: TSS, exon IGR, 5'UTR, 3'UTR (b, d).

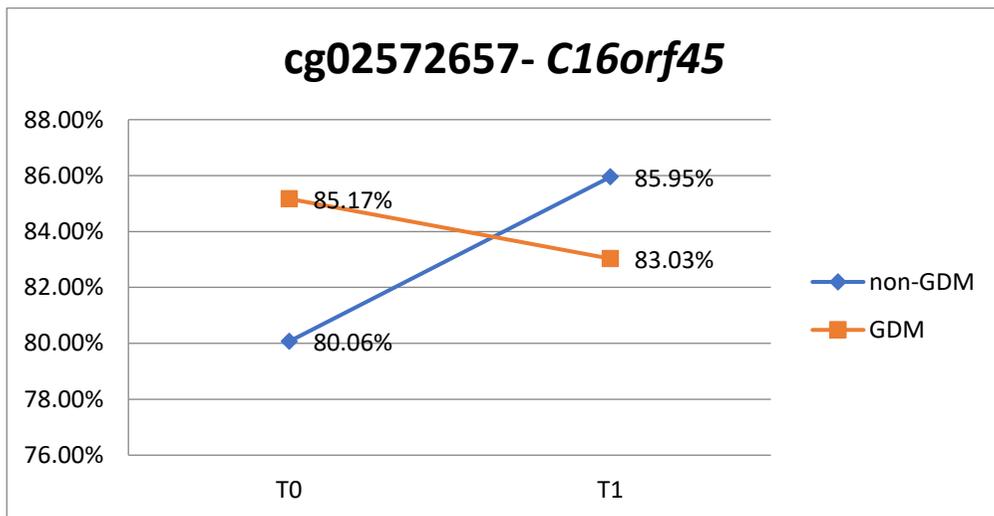
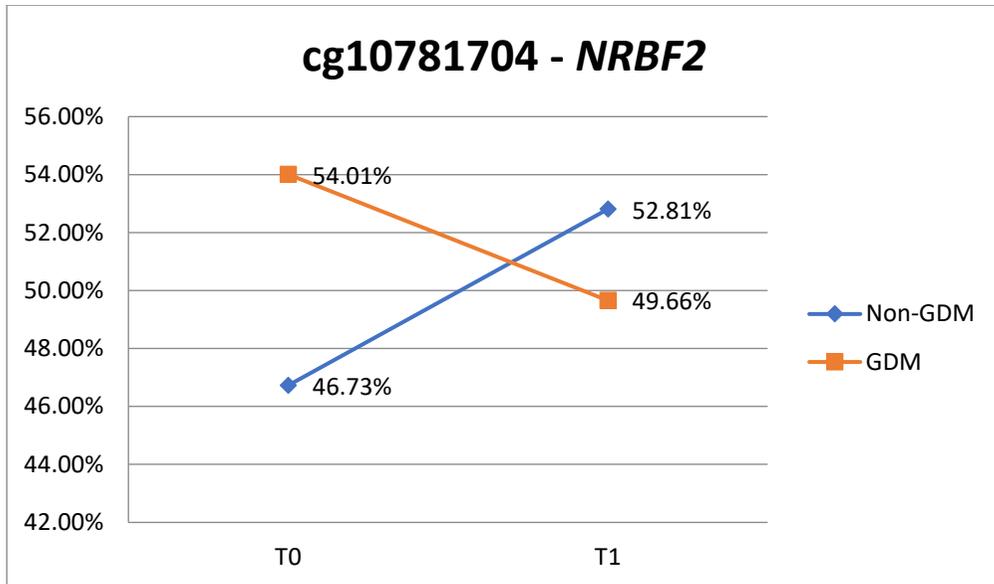
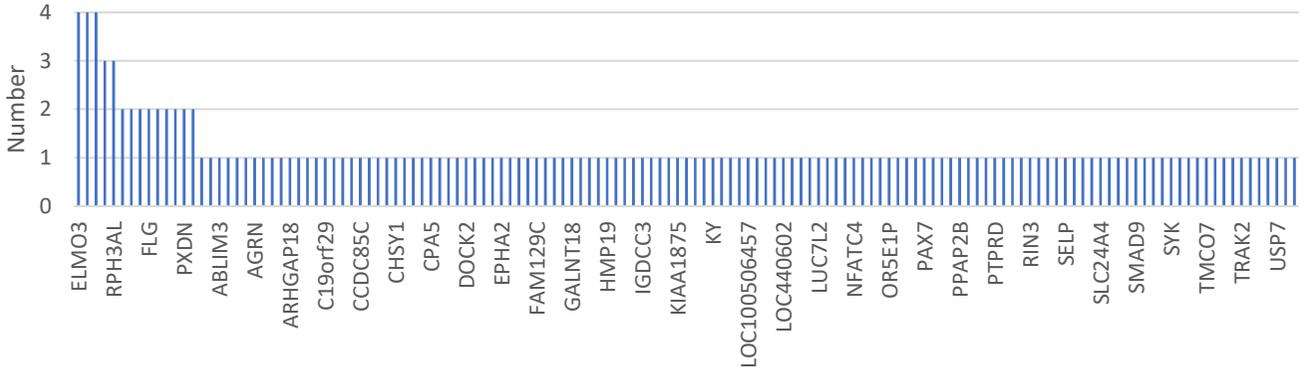


Figura suplementaria 3- Niveles de metilación de dos CpGs que responden diferencialmente a lo largo del tiempo en el grupo de DG relativamente al grupo de no-DG en la visita inicial (T0) y preparto (T1).

A) Top 100 Gene Mostly Enriched by Significant CpG



B) Top 14 Gene Mostly Enriched by Significant CpG

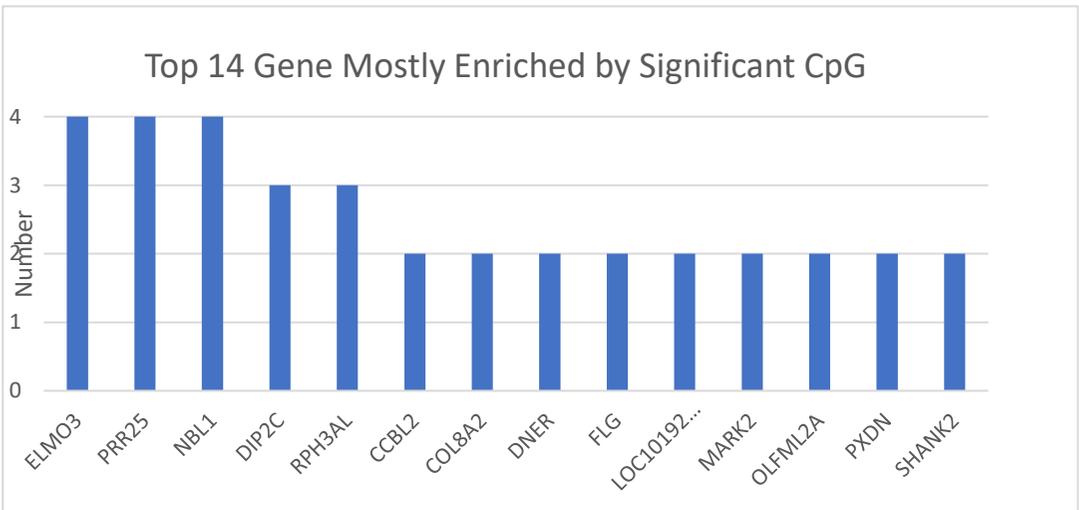


Figura suplementaria 4- Bar plot de los genes más enriquecidos de los DMPs significativos. El número de cada barra indica el número de CpGs que están incluidos en ese gene.

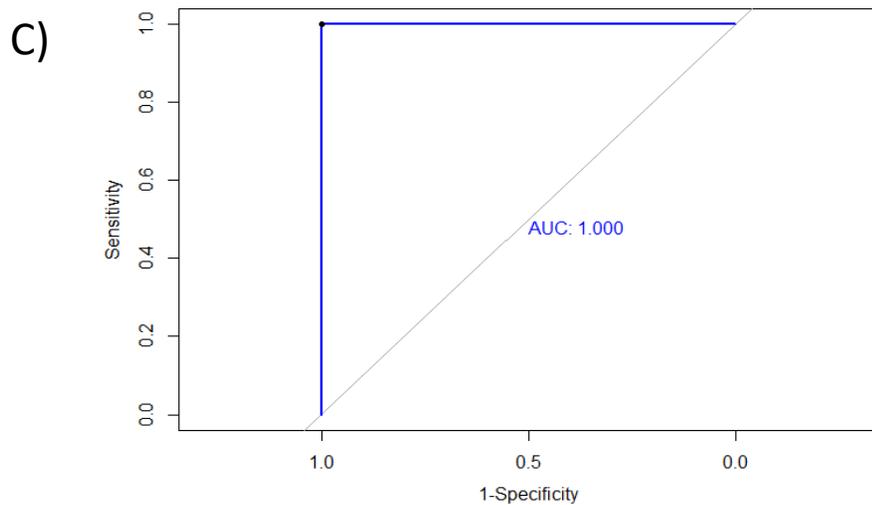
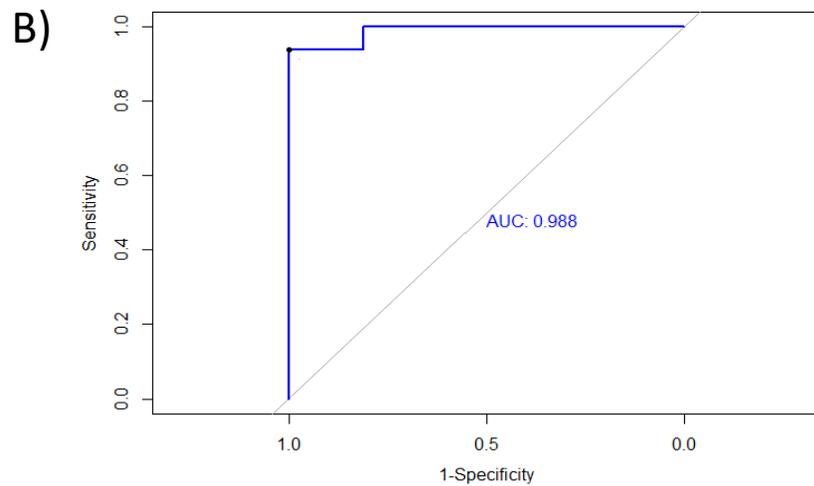
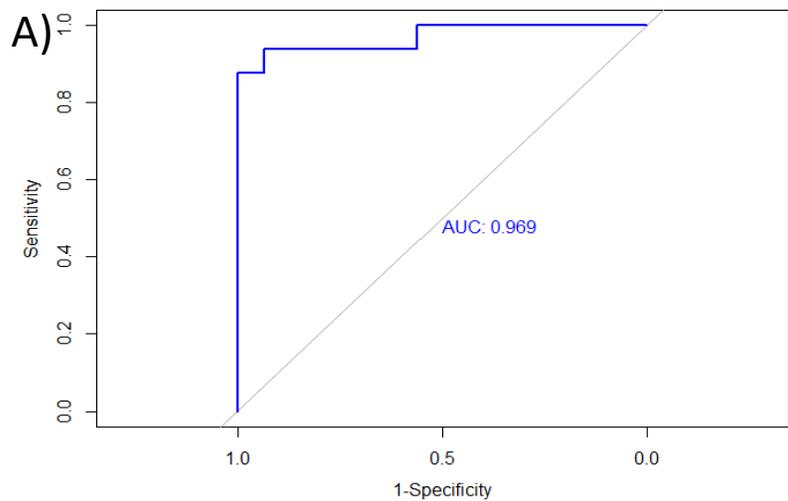


Figura suplementaria 5- Modelos de curvas ROC de los tres mejores CpG que discriminan entre DG y no DG: A) Modelo con un solo CpG cg01459453. B) Modelo con dos CpG: cg01459453 y cg15329406. C) Modelo con 3 CpG: cg01459453, cg15329406 y cg04095097. AUC: área bajo la curva.

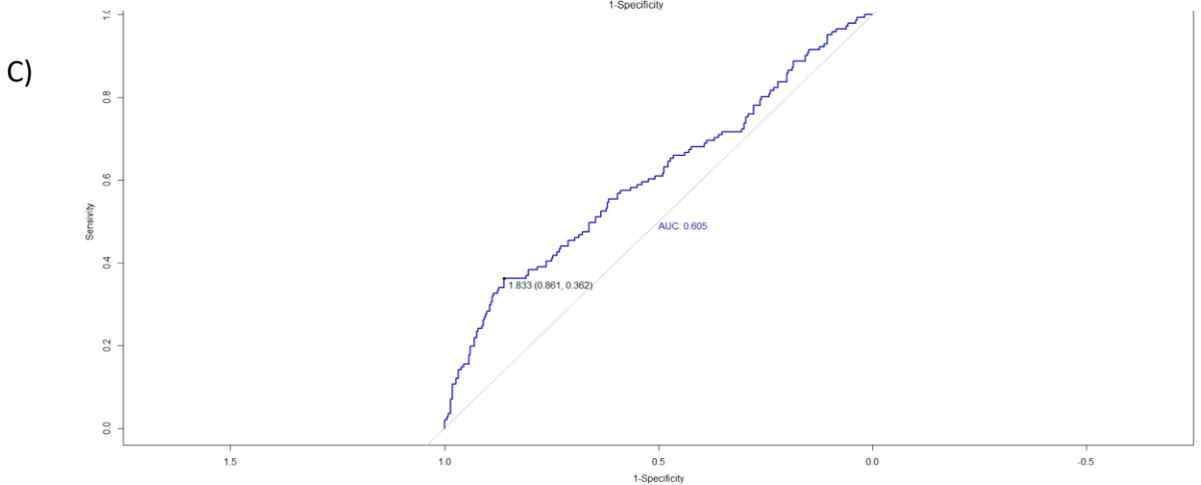
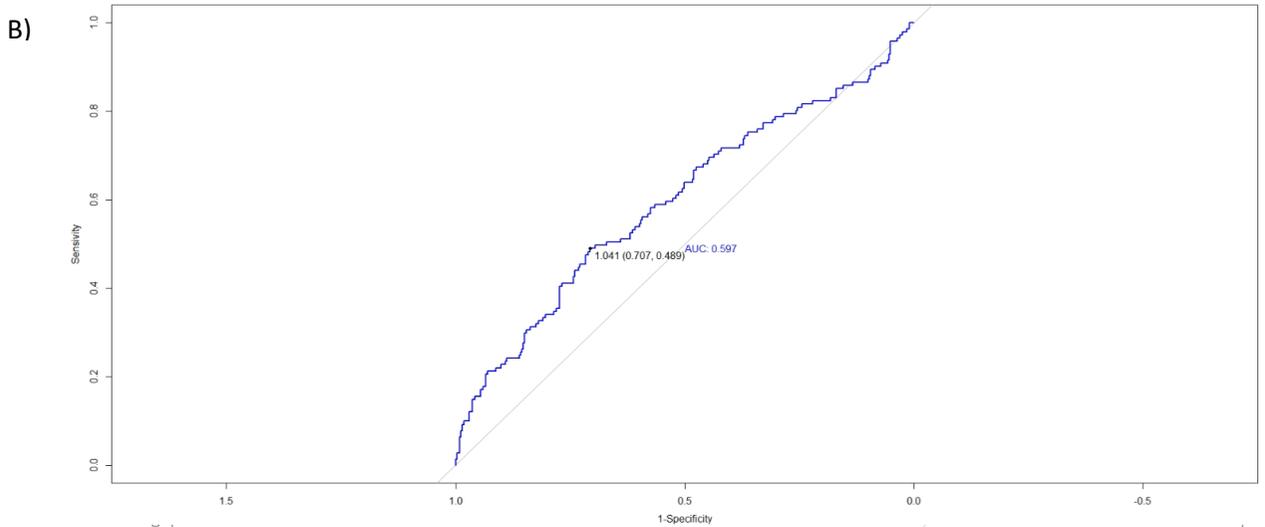
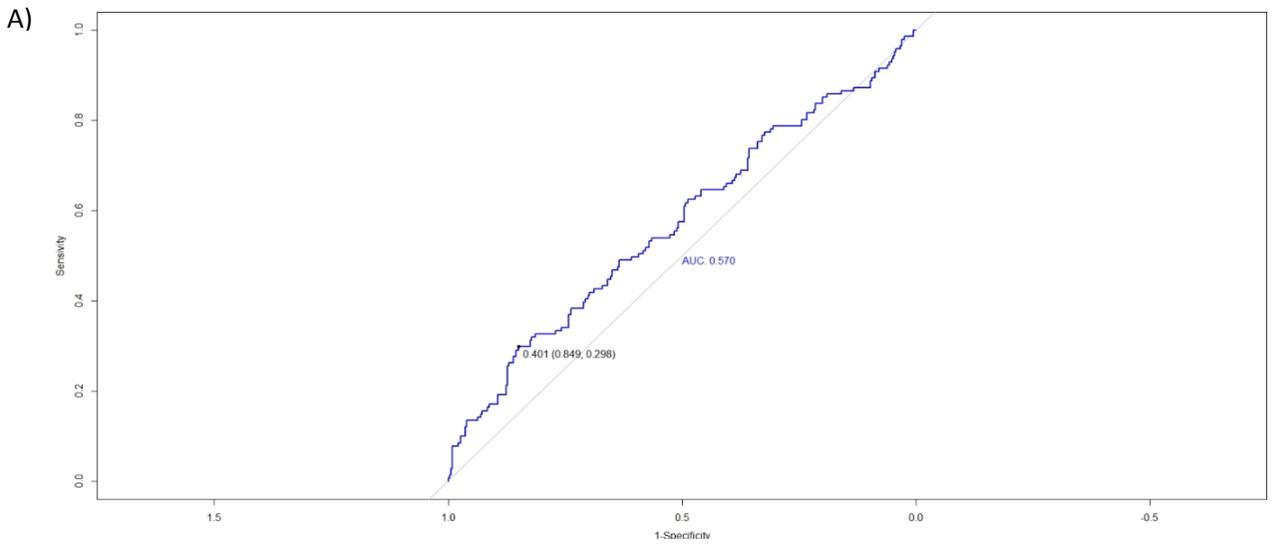
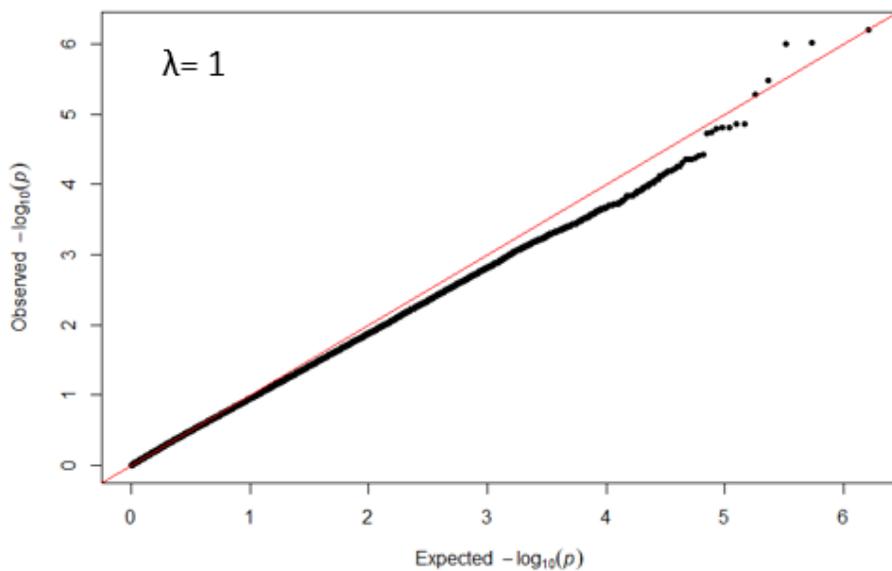


Figura suplementaria 6- Modelos de curvas ROC de los 3 CpG validados en la cohorte EPIREG. A) Modelo de un CpG: cg04095097, B) Modelo de dos CpG: cg04095097, cg27603605, C) Modelo de tres CpG: cg04095097, cg27603605 y cg12080079. AUC: área bajo la curva.

A) Posparto



B) Año

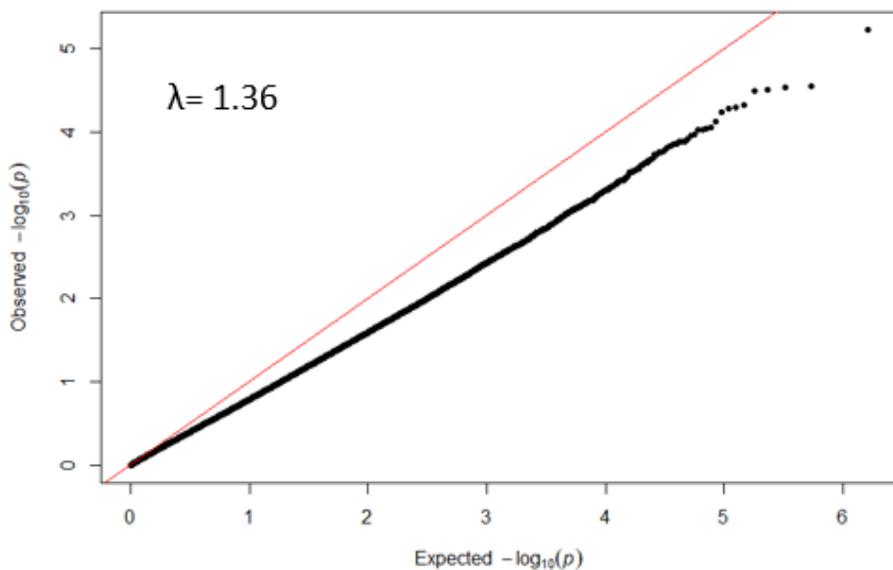


Figura suplementaria 7- QQplot del EWAS de saliva de la descendencia A) Visita posparto λ 1 no hay inflación y hay baja probabilidad de que haya falsos positivos B) visita al año con una λ de 1.36 que se considera moderada, por hay más probabilidad de encontrarnos con falsos positivos.

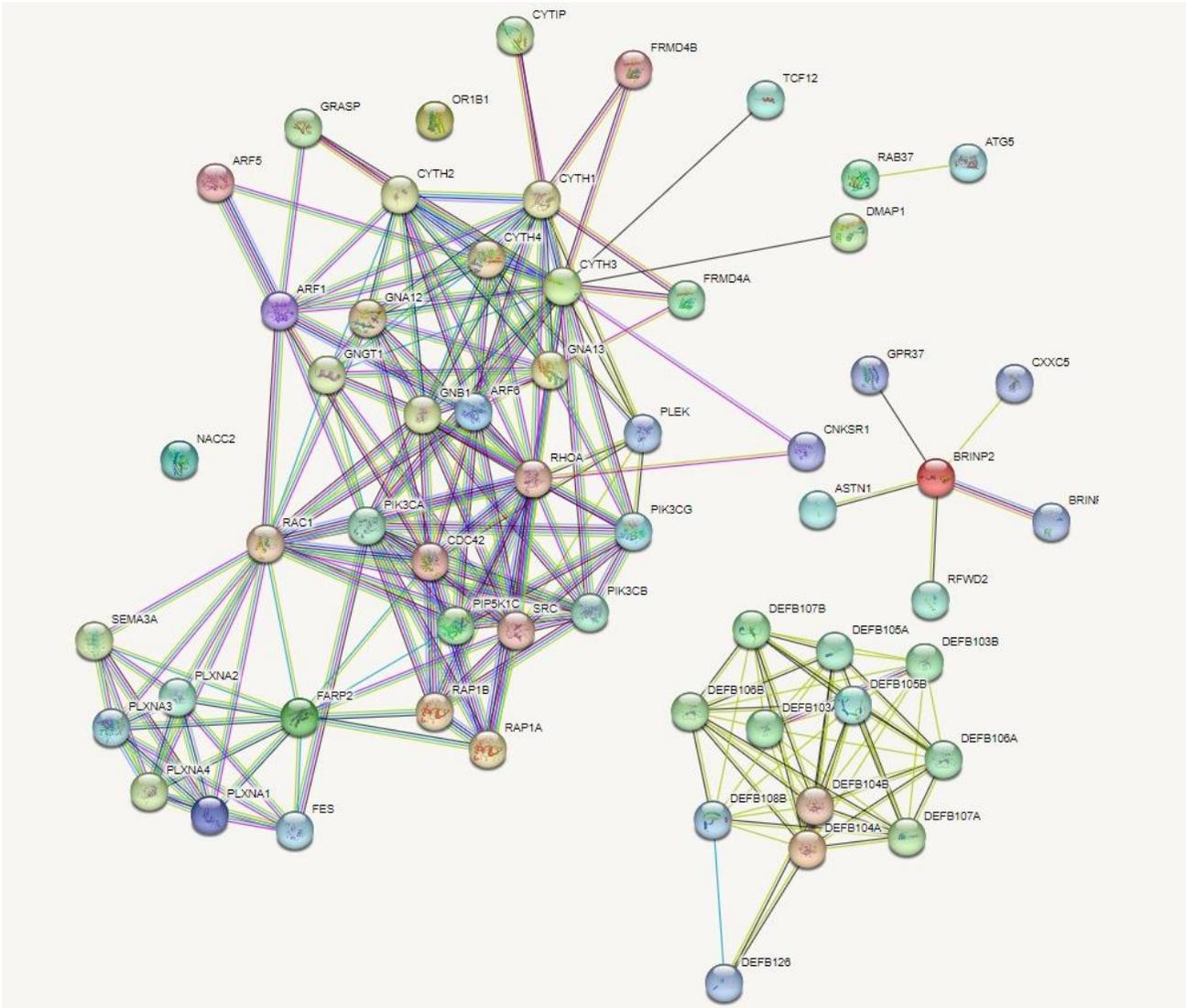


Figura suplementaria 8- Red de STRING de interacciones proteicas resultante de introducir las proteínas en los que se encuentran los 16 CpGs en común y haber configurado STRING para obtener con qué otras proteínas interaccionan.

Abreviaturas

ABCG1: Miembro 1 de la subfamilia G de casetes de unión a ATP/ATP Binding Cassette Subfamily G Member 1

ADA: American Diabetes Association

ADCYAP1R1: Receptor tipo I de ADCYAP/ADCYAP Receptor Type I

AGRN: Agrin

AKT2: AKT serina/treonina quinasa 2/ AKT Serine/Threonine Kinase 2

AUC: Área Bajo la Curva/Area Under the Curve

BMP: Proteína morfogenética ósea/Bone morphogenetic protein

BRINP2: BMP/Ácido retinoico inducible específico neural 2 /BMP/Retinoic Acid Inducible Neural Specific 2

C5orf34: Marco de lectura abierta 34 del cromosoma 5/Chromosome 5 Open Reading Frame 34

CAMTA1: Activador de la Transcripción de unió a calmodio 1/Calmodium Binding Transcription Activator 1

CC: Carpenter y Coustan

CCDC124: Dominio Coiled-Coil que contiene 124/Coiled-Coil Domain Containing 124

CDH18: Cadherin 18

CDKAL1: Proteína Asociada a la Subunidad Reguladora CDK5 1 Like 1/CDK5 Regulatory Subunit Associated Protein 1 Like 1

CHSY1: Condrotín sulfato sintasa 1/ Chondroitin Sulfate Synthase 1

COPS8: Subunidad 8 del Singalosoma COP9/COP9 Signalosome Subunit 8

CTIF: Factor de iniciación de la traducción dependiente de complejo de unión a la capsula/ Cap Binding Complex Dependent Translation Initiation Factor

CYTH3: Cytohesina 3/ Cytohesin 3

DEFB104A: Defensina Beta 104A/ Defensin Beta 104A

DIP2C: Proteína 2 homóloga de interacción con el disco C /Disco Interacting Protein 2 Homolog C

DMG/GDM: Diabetes Mellitus Gestacional/Gestational Diabetes Mellitus

DMP: Posición Diferencialmente Metilada/ Differentially Methylated Position

DMR: Región Diferencialmente Metilada/ Differentially Methylated Region

DMT1: Diabetes mellitus tipo 1

DMT2/TMD2: Diabetes mellitus tipo 2/Type 2 Diabetes Mellitus

DOHAD: Orígenes evolutivos de la salud y la enfermedad/Developmental Origins of Health and Disease

EG: Edad Gestacional

EGFR:

ELISA: Kit de Inmunoensayo Enzimático

EM-seq: Metilación por secuenciación Enzimática/Enzymatic Methylation sequencing

ErbB: Receptor Tirosina Quinasa 2 Erb-B2/Erb-B2 Receptor Tyrosine Kinase 2

ERBB4: Receptor Erb-B2 tirosina quinasa 4/Erb-B2 Receptor Tyrosine Kinase 4

EWAS: Estudio de asociación de todo el epigenoma/Epigenome-wide Association Study

FAMB5B: BMP/Ácido retinoico inducible específico neural 2/BMP/Retinoic Acid Inducible Neural Specific 2

FARP2: FERM, ARH/RhoGEF y proteína 2 de dominio Pleckstrin /FERM, ARH/RhoGEF And Pleckstrin Domain Protein 2

FDR: False Discovery Ratio

GCK: Glucoquinasa/Glucokinase

GEDE: Grupo Español de Diabetes y Embarazo

Gluc0: Glucosa en ayunas

Gluc2: Glucosa 2h tras la sobre carga oral

GLUT4: Transportador de glicosa 4

GO: Ontología de genes/Gene Ontology

GRS: Puntuación de riesgo genético/ Genetic Risk Score

GWAS: Estudio de asociación del genoma completo/ Genome-wide association study

HAAO: 3-Hidroxiantranilato 3,4-Dioxigenasa/ 3-Hydroxyanthranilate 3,4-Dioxygenase

HAPO: Hiperglucemia y efectos adversos durante el embarazo/ Hyperglycemia and Adverse Pregnancy Outcomes

HAPO-FUS: Seguimiento del estudio Hiperglucemia y efectos adversos durante el embarazo/ Hyperglycemia and adverse pregnancy Outcome Follow-Up Study

HbA1C: Hemoglobina A1C glucosilada

HDL: Lipoproteínas de Alta Densidad/ High Density Lipoprotein

HOMA-IR: Índice de Resistencia a la insulina según el modelo de Homeostasis/ Homeostasis Model Assesment of Insulin Resistance Index

HPL: Hormonas Liberadas por la placeta

HSD17B11: Hidroxiesteroide 17-Beta Deshidrogenasa 11/ Hydroxysteroid 17-Beta Dehydrogenase 11

IADPSG: Asociación Internacional del Grupo de Estudios sobre Diabetes y Embarazo/The International Association of the Diabetes and Pregnancy Study Group

IDAT: Intensity Data

IDF: Federación Internacional de Diabetes/ International Diabetes Federation

IGF-1: Factor de crecimiento 1 similar a la insulina/Insulin Like Growth Factor 1

IGF2BP2: Factor de crecimiento similar a la insulina 2 Proteína de unión al ARNm 2 /Insulin Like Growth Factor 2 MRNA Binding Protein 2

IMC: índice de masa corporal

INS: Insulina/ Insulin

IRS1: Sustrato 1 del receptor de la insulina

IR β : Subunidad β de la insulina

KCNJ11: Miembro de la Subfamilia J de los canales de potasio rectificadores internos/Potassium Inwardly Rectifying Channel Subfamily J Member

KCNQ1: Canal dependiente de Potasio cana de la subfamilia Q miembro 1/Potassium Voltage-Gated Channel Subfamily Q Member 1

KEGG: Enciclopedia de Genes y Genomas de Kioto/ Kyoto Encyclopedia of Genes and Genomes

LEF1: Factor 1 de unión a potenciadores linfoides /Lymphoid Enhancer Binding Factor 1

LGAL3BP: Proteína de union a galectina 3/Galectin 3 Binding Protein

LGR6: Receptor 6 acoplado a proteína G que contiene repeticiones ricas en leucina/ Leucine Rich Repeat Containing G Protein-Coupled Receptor 6

LINC00917: ARN intergénico largo no codificante de proteínas 917/Long Intergenic Non-Protein Coding RNA 917

MAF: Frecuencia de alelo menos común/ Minor Allele Frequency

MAP4K2: Proteína quinasa avtivada por mitógenos quinasa quinasa quinasa quinasa 2/Mitogen-Activated Protein Kinase Kinase Kinase Kinase 2

MAPK:

MeDIP: Inmunoprecipitación del ADN metilado/ Methylated DNA Immunoprecipitation

MINOS1-NBL1: MICOS10-NBL1 Lectura/ MICOS10-NBL1 Readthrough

mQTL: Loci de riesgo cuantitativo de metilación/ Methylation Quantitative Trait Loci

MRS: Puntuación de riesgo de metilación/ Methylation Risk Score

MS-MLPA: Amplificación de sonda dependiente de ligadura múltiple específica de metilo/Methylation Specific Multiplex Ligation-Dependent Probe Amplification

MSRE: Digestión con enzimas de restricción sensibles a la metilación/ Methylation-sensitive restriction enzume and multiplexed

MTNR1B: Receptor de la Melatonina 1B/ Melatonin Receptor 1B

MTOR: Diana mecánica de la rapamicina quinasa/Mechanistic Target Of Rapamycin Kinase

NBL1: Antagonista de BMP de la familia DAN/DAN Family BMP Antagonist

Nc-RNA: ARN no codificante/ non-coding ARN

NDDG: Grupo Nacional de datos sobre Diabetes /National Diabetes Data Group

OMS: Organización Mundial de la Salud

ONT: Oxford Nanopore Technologies

OR: Odd Ratio

PacBio: PacificBiosciences

PACE: Epigenética del embarazo y la infancia/Pregnancy and Childhood Epigenetics

PAS: Presión Arterial Sistólica

PCR: Reacción en Cadena de la Polimerasa/ Polymerase Chain Reaction

PDIA5: Proteína Disulfuro Isomerasa Familia A Miembro 5/Protein Disulfide Isomerase Family A Member 5

PGH: Hormonas de Crecimiento placentario

PI3K: Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Delta

PIK3R5: Subunidad reguladora 5 de la fosfoinositida-3-quinasa/ Phosphoinositide-3-Kinase Regulatory Subunit 5

PPI: Interacción Proteína-Proteína/ Protein-Protein Interaccion

PRKAA2: Subunidad catalítica alfa 2 de la proteína cinasa activada por AMP/Protein Kinase AMP-Activated Catalytic Subunit Alpha 2

PRKAA2: Subunidad catalítica alfa 2 de la proteína quinasa activada por AMP/ Protein Kinase AMP-Activated Catalytic Subunit Alpha 2

PTGIS: Prostaglandina Sintasa/Prostaglandin I2 Synthase

PTPRN2: Receptor de proteína tirosina fosfatasa tipo N2/Protein Tyrosine Phosphatase Receptor Type N2

PTPRN2: SCY1 Como Pseudoquinasa 3/SCY1 Like Pseudokinase 3

QTL: Loci de riesgo cuantitativo/ Quantitative Trait Loci

ROC: Característica Operativa del Receptor/Receiver Operating Characteristic

RPH3AL: Rabfilina similar a 3A /Rabphilin 3A Like

RPS6KA2: Proteína ribosómica S6 quinasa A2/Ribosomal Protein S6 Kinase A2

RRBS: Secuenciación por bisulfito de representación reducida/Reduced Representation Bisulfite Sequencing

SELP: Selectina B/Selectin B

SLC9A3: Portador de solute familia 9 miembro A3/Solute Carrier Family 9 Member A3

SLC9A3R1/ NHERF: Proteína de andamiaje PDZ 1 de la familia NHERF/NHERF Family PDZ Scaffold Protein 1

SMRT: Secuenciación a tiempo real

SNP: Polimorfismo de nucleótido único/Single Nucleotide Polimorfims

SOG/OGTT: Sobrecarga Oral de la Glucosa/Oral Glucose Tolerance Test

SVD: Descomposición de los valores singulares/ Singular Value Decomposition

SYNGR1: Sinaptogirina 1/Synaptogyrin 1

SYNPR: Synaptoporin

T.0: Visita inicial

T.1: Visita parto

T.2: Visita posparto

T.3: Visita al año

TCF7L2: Factor de transcripción 7 similar a 2/Transcription Factor 7 Like 2

TG: Triglicéridos

TGF- β : Factor de crecimiento transformante Beta 1/Transforming Growth Factor Beta 1

TNF: Factor de necrosis tumoral/ Tumor Necrosis Factor

TRAPPC9: Subunidad 9 del complejo proteico de partículas de tráfico/Trafficking Protein Particle Complex Subunit 9

TRIP6: Interactor 6 del receptor de la hormona tiroidea/Thyroid Hormone Receptor
Interactor 6

TSS: Sitio de inicio de la transcripción/ Transcription Start Site

UTR: Región no traducida/ Untranslated Region

VEGF: Factor A de crecimiento endotelial vascular/Vascular Endothelial Growth Factor A

