



**ASOCIACIÓN DE GENES DE LA CADENA MITOCONDRIAL
RESPIRATORIA EN EL DESARROLLO DE DIABETES
MELLITUS TIPO 2 Y PARAMETROS RELACIONADOS**

Tesis Doctoral

Presentada por:

Griselda De Marco

Dirigida por:

F. Javier Chaves Martínez

Rafael Carmena Rodríguez

PROGRAMA DE DOCTORADO EN BIOQUÍMICA Y BIOMEDICINA

Departamento de Bioquímica y Biología Molecular

UNIVERSITAT DE VALÈNCIA

Valencia, mayo 2017



VNIVERSITAT
DE VALÈNCIA

INCLIVA | VLC
Instituto de Investigación Sanitaria

D. Felipe Javier Chaves Martínez, doctor en Ciencias Biológicas, investigador y director de la Unidad de Genómica y Diagnóstico Genético en la Fundación de Investigación del Hospital Clínico de Valencia-INCLIVA, y D. Rafael Carmena Rodríguez, Catedrático Emérito de la Unidad Docente de Medicina de la Universitat de Valencia, informan:

Que Doña Griselda O. De Marco Romero, Licenciada en Bioquímica por la Universidad de La Plata, Buenos Aires, ha realizado bajo nuestra dirección en la Fundación de Investigación del Hospital Clínico de Valencia-INCLIVA, en la Unidad de Genómica y Diagnóstico Genético, el trabajo experimental conducente a la elaboración de la presente memoria de investigación, destinada a Tesis Doctoral, que se presenta bajo el título de:

“ASOCIACIÓN DE GENES DE LA CADENA MITOCONDRIAL RESPIRATORIA EN EL DESARROLLO DE DIABETES TIPO 2 Y PARAMETROS RELACIONADOS”

Y para que así conste, se expide el presente, en Valencia a 26 de mayo de 2017.

Fdo. Dr. F. Javier Chaves Martínez Fdo. Dr. Prof. Rafael Carmena Rodríguez

Este trabajo de tesis doctoral ha sido financiado por el proyecto coordinado “Cooperative population and database studies for genetic association analysis in Type 2 Diabetes mellitus and related traits” (INGENFRED, CIBER-02-08-2009) del Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas asociadas (CIBERDEM), cuyo investigador principal es el Dr. Prof. Rafael Carmena Rodríguez, y el coordinador es el Dr. F. Javier Chaves Martínez.

En memoria de mi padre, adiós querido papá

Agradecimientos

La realización de una tesis conlleva mucho trabajo y esfuerzo, pero más aún, requiere de un apoyo constante de la gente que comparte en alguna medida este camino. En primer lugar, quiero agradecer la oportunidad que me brindaron mis directores de tesis, el Dr. Rafael Carmena y el Dr. Javier Chaves, de poder incorporarme a la Unidad de Genómica y Diagnóstico Genético, para llevar a cabo este trabajo, y aprender tantas cosas en estos años. Todos los integrantes de la UGDG, me han acompañado en este proyecto, empezando por mi “maestra”, la que me enseñó las bases de cómo manejarme en un laboratorio en el día a día, siempre con una sonrisa, y mucha paciencia, gracias María Luisa. A todos les debo el poder sentirme como en casa, estando tan lejos de ella. Gracias Tani, Victoria, Cristina, YoVero, Vero Navarro, Kike, Laura, Pilar, Veronik, Inma, Vane, Ana Bárbara, Raquel, Javi, Pablo, Azahara, Ana. Les agradezco que me hayan apoyado, y que hayan confiado en mí, siempre viendo el lado positivo de las cosas, y ayudándome en todo, en lo laboral, pero también en otros aspectos de mi vida.

Agradezco a la Unidad de Bioinformática del Incliva, especialmente a Azahara, por haberme ayudado en el análisis de resultados de secuenciación masiva de la tesis.

En lo personal, creo que si no fuera por mi padre, no estaría aquí, investigando sobre la diabetes...Gracias papá por darme esta vocación, y es un dolor muy grande para mí que al final no estés aquí ya conmigo para verlo. Gracias mamá, por estar siempre en los malos momentos. Les agradezco a toda mi familia, mi hermana Faby y mi hermano Rubén, a mis sobrinos Augusto, Milena, Cloe, Simón y Raquel, que hayan aceptado el estar lejos, pero sintiéndonos siempre cerca. Y a mis mejores amigos, Lety, Manu, Laura, Lorena, Ali, Meche, Sil, Silvina, Luciano, Laurita, gracias por hacerme sentir querida en la distancia. Y no puedo dejar de decir unas palabras de mi compañero incondicional de tantos años, Daniel, que fue el que me empujó a estar aquí, y me acompañó en todo este camino, dándome además ese precioso “regalo” llamado Ariadna.

Abreviaturas

| | |
|----------|-------------------------------------------------------------------------|
| ADA | American Diabetes Association |
| ADN | Acido desoxirribonucleico |
| ADP | Adenosina difosfato |
| ARN | Acido ribonucleico |
| ATP | Adenosina trifosfato |
| ATP III | Adult Treatment Panel III |
| CMR | Cadena mitocondrial respiratoria |
| COX7C | Subunidad 7C de la citocromo c oxidasa DG |
| DM | Diabetes Mellitus |
| DMT1 | Diabetes Mellitus tipo 1 |
| DMT2 | Diabetes Mellitus tipo 2 |
| EASD | European Association for the Study of Diabetes |
| EMBL-EBI | European Molecular Biology Laboratory-European Bioinformatics Institute |
| GAA | Glucosa alterada en ayunas |
| GWAS | Estudios de asociación de genoma completo |
| HbA1C | Hemoglobina glicosilada A1C |
| HTA | Hipertensión arterial |
| HWE | Equilibrio de Hardy-Weinberg |
| IDF | Internacional Diabetes Federation |
| IMC | Indice de masa corporal |
| LADA | Diabetes latente autoinmune del adulto |
| MAF | Frecuencia del alelo minoritario |
| MODY | Diabetes de la juventud de aparición en la madurez |
| mtADN | ADN mitocondrial |
| NCBI | National Center of Biotechnology Information |
| NDUFS2 | Subunidad S2 del núcleo de la NADH ubiquinona oxidoreductasa |
| NDUFS3 | Subunidad S3 del núcleo de la NADH ubiquinona oxidoreductasa |
| NDUFS5 | Subunidad S5 del núcleo de la NADH ubiquinona oxidoreductasa |
| NADH | nicotinamida adenina dinucleótido forma reducida |
| NGS | Secuenciación de nueva generación |
| OB | Obesidad |

| | |
|------------|--------------------------------------------------------|
| OB Central | Obesidad Central |
| OMS | Organización mundial de la salud |
| OXPPOS | fosforilación oxidativa |
| PAD | Presión arterial diastólica |
| PAS | Presión arterial sistólica |
| RI | Resistencia a la insulina |
| ROS | Especies reactivas del oxígeno |
| SNP | Polimorfismo de un solo nucleótido |
| TAG | Tolerancia alterada a la glucosa |
| UQCRB | Proteína de unión a la ubiquinol-citocromo c reductasa |

Indice

| | |
|------------------------------------------------------------------------------------------|----|
| Resumen | 1 |
| 1 Introducción | 3 |
| 1.1 Diabetes mellitus | 3 |
| 1.1.1 Descripción general y clasificaciones | 3 |
| 1.1.2 Criterios diagnósticos de la diabetes | 8 |
| 1.2 Diabetes mellitus tipo 2 | 9 |
| 1.2.1 Prevalencia de la diabetes mellitus tipo 2 | 9 |
| 1.2.2 Fisiopatología de la DMT2 | 10 |
| 1.2.3 Factores que afectan el desarrollo de la DMT2 | 12 |
| 1.2.3.1 Factores genéticos de la DMT2 | 13 |
| 1.2.3.2 Factores de riesgo de la DMT2 | 21 |
| 1.2.3.2.1 Obesidad | 23 |
| 1.2.3.2.2 Obesidad central | 27 |
| 1.2.3.2.3 Hipertensión arterial | 28 |
| 1.2.4 Tratamiento DMT2 | 32 |
| 1.3 Cadena mitocondrial respiratoria (CMR) | 36 |
| 1.3.1 Características generales y mecanismo molecular | 36 |
| 1.3.2 Cadena mitocondrial respiratoria y su relación con la DMT2 | 40 |
| 2 Hipótesis | 45 |
| 3 Objetivos | 46 |
| 4 Metodología | 47 |
| 4.1 Diseño de la tesis | 47 |
| 4.2 Descripción de las poblaciones analizadas | 48 |
| 4.2.1 Estudio Hortega | 48 |
| 4.2.2 Estudio Pizarra | 49 |
| 4.2.3 Estudio Segovia | 49 |
| 4.2.4 Estudio Di@bet.es | 50 |
| 4.3 Medidas bioquímicas | 51 |
| 4.4 Fase 1: Estudio de los polimorfismos de la CMR | 51 |
| 4.4.1 Extracción de ADN a partir de sangre periférica | 51 |
| 4.4.2 Selección de las variantes genéticas | 52 |
| 4.4.3 Método de genotipado de los polimorfismos de la CMR | 55 |
| 4.5 Fase 2: Identificación de variaciones genéticas en los genes asociados con DMT2 y OB | 63 |

| | |
|--------------------------------------------------------------------------------------------|-----|
| 4.6 Fase3: Estudio de las variaciones identificadas en la fase 2 | 69 |
| 4.6.1 Preparación de las librerías para la secuenciación masiva | 71 |
| 4.6.2 Secuenciación de las muestras en el sistema Miseq | 75 |
| 4.6.3 Análisis bioinformático de los resultados | 82 |
| 4.6.3.1 Análisis primario de los resultados de la secuenciación de nueva generación | 82 |
| 4.6.3.2 Análisis secundario de los resultados de la secuenciación de nueva generación | 87 |
| 4.7 Fase 4: Estudios de replicación de las asociaciones halladas | 90 |
| 4.8 Análisis estadísticos | 95 |
| 4.8.1 Estadística descriptiva | 95 |
| 4.8.2 Estudios de asociación | 96 |
| 5 Resultados | 99 |
| <i>5.1 Fase 1: Estudio de polimorfismos de la CMR</i> | 99 |
| 5.1.1 Características generales de las poblaciones analizadas | 99 |
| 5.1.2 Asociaciones de los polimorfismos de la CMR con la OB, DMT2 Y HTA | 99 |
| 5.1.2.1 Asociaciones de los polimorfismos de la CMR y la OB | 100 |
| 5.1.2.1.1 Resultados de asociaciones de SNPs de CMR y OB en las poblaciones individuales | 100 |
| 5.1.2.1.2 Resultados de asociaciones de SNPs de CMR y OB en la población total | 103 |
| 5.1.2.1.3 Resultados de asociaciones de SNPs de CMR y OB en subpoblaciones | 104 |
| 5.1.2.1.3.1 Resultados de asociaciones según género | 105 |
| 5.1.2.2 Asociaciones de los polimorfismos de la CMR y la DMT2 | 106 |
| 5.1.2.2.1 Resultados de asociaciones de SNPs de CMR y DMT2 en las poblaciones individuales | 106 |
| 5.1.2.2.2 Resultados de asociaciones de SNPs de CMR y DMT2 en la población total | 108 |
| 5.1.2.2.3 Resultados de asociaciones de SNPs de CMR y DMT2 en subpoblaciones | 111 |
| 5.1.2.2.3.1 Resultados de asociaciones según género | 111 |
| 5.1.2.2.3.2 Resultados de las asociaciones en población obesa | 113 |
| 5.1.2.2.3.3 Resultados de las asociaciones en población no obesa | 114 |
| 5.1.2.3 Asociaciones de los polimorfismos de la CMR y la HTA | 115 |
| 5.1.2.3.1 Resultados de asociaciones de SNPs de CMR y HTA en las poblaciones individuales | 115 |
| 5.1.2.3.2 Resultados de asociaciones de SNPs de CMR y HTA en la población total | 117 |
| 5.1.2.3.3 Resultados de asociaciones de SNPs de CMR y HTA en subpoblaciones | 118 |
| 5.1.2.3.3.1 Resultados de asociaciones según género | 118 |
| 5.1.2.3.3.2 Resultados de las asociaciones en población obesa | 120 |
| 5.1.2.3.3.3 Resultados de las asociaciones en población no obesa | 121 |
| <i>5.2 Fase 2: Identificación de variaciones genéticas en los genes asociados</i> | 122 |

| | |
|------------------------------------------------------------------------------------------------|-----|
| 5.3 Fase 3: Estudio de las variaciones identificadas en la fase 2 | 124 |
| 5.3.1 Control de calidad de la secuenciación | 125 |
| 5.3.2 Cobertura de los amplicones secuenciados | 129 |
| 5.3.3 Asociaciones de los polimorfismos seleccionados en la fase 2 con DMT2, OB y HTA | 129 |
| 5.3.3.1 Asociaciones con OB de los polimorfismos seleccionados, en la población de la fase 1 | 129 |
| 5.3.3.2 Asociaciones con DMT2 de los polimorfismos seleccionados, en la población de la fase 1 | 131 |
| 5.3.3.3 Asociaciones con HTA de los polimorfismos seleccionados, en la población de la fase 1 | 132 |
| 5.4 Fase 4: Estudios de replicación de las asociaciones halladas | 133 |
| 5.4.1 Características generales de la población analizada | 133 |
| 5.4.2 Control de calidad de la secuenciación | 134 |
| 5.4.3 Cobertura de los amplicones secuenciados | 138 |
| 5.4.4 Asociaciones de los polimorfismos de la CMR y la OB en la población Di@bet.es | 138 |
| 5.4.4.1 Resultados de asociaciones de SNPs de CMR y OB en subpoblaciones | 139 |
| 5.4.4.1.1 Resultados de las asociaciones de CMR Y OB según género | 139 |
| 5.4.5 Asociaciones de los polimorfismos de la CMR y la DMT2 en la población Di@bet.es | 141 |
| 5.4.5.1 Resultados de asociaciones de SNPs de CMR y DMT2 en subpoblaciones | 142 |
| 5.4.5.1.1 Resultados de las asociaciones de CMR Y DMT2 según género | 142 |
| 5.4.5.1.2 Resultados de las asociaciones de CMR Y DMT2 en individuos obesos | 144 |
| 5.4.5.1.3 Resultados de las asociaciones de CMR Y DMT2 en individuos no obesos | 145 |
| 5.4.6 Asociaciones de los polimorfismos de la CMR y la HTA en la población Di@bet.es | 146 |
| 5.4.6.1 Resultados de asociaciones de SNPs de CMR y HTA en subpoblaciones | 147 |
| 5.4.6.1.1 Resultados de las asociaciones de CMR Y HTA según género | 147 |
| 5.4.6.1.2 Resultados de las asociaciones de CMR Y HTA en individuos obesos | 148 |
| 5.4.6.1.3 Resultados de las asociaciones de CMR Y HTA en individuos no obesos | 149 |
| 5.4.7 Validación de las asociaciones replicadas en las poblaciones | 150 |
| 5.4.7.1 Asociaciones replicadas en las poblaciones generales totales | 150 |
| 5.4.7.2 Comparación de las asociaciones con OB obtenidas en la fase 1 y 4 | 150 |
| 5.4.7.3 Comparación de las asociaciones con DMT2 obtenidas en la fase 1 y 4 | 151 |
| 5.4.7.4 Comparación de las asociaciones con HTA obtenidas en la fase 1 y 4 | 153 |
| 5.4.7.5 Comparación de las asociaciones con OB, DMT2 y HTA obtenidas en la fase 3 y 4 | 157 |
| 6. Discusión | 159 |
| 7. Conclusiones | 178 |
| 8 Bibliografía | 179 |
| 9 Anexo | 199 |

Resumen

Introducción: la diabetes mellitus tipo 2 (DMT2) es una enfermedad compleja, resultado de la combinación de factores ambientales, genéticos y adquiridos. Se ha convertido en una epidemia global, que contribuye significativamente a la morbi-mortalidad prematura. Su heredabilidad es de un 70%; no obstante, los estudios realizados sólo han podido identificar una pequeña proporción de este componente genético. La cadena mitocondrial respiratoria (CMR) es esencial en el metabolismo energético y en la producción de especies reactivas del oxígeno (ROS). Un mal funcionamiento de este sistema biológico parece estar implicado en la DMT2 y otras alteraciones relacionadas (obesidad, hipertensión). Sin embargo, sólo se han estudiado hasta el momento unos pocos genes de los que codifican proteínas de la CMR.

Objetivos: estudiar la posible asociación entre genes pertenecientes a la CMR con el riesgo de padecer DMT2, parámetros relacionados, y otros desórdenes asociados, como la obesidad y la hipertensión.

Metodología:

Fase 1: análisis de 96 polimorfismos localizados en 29 genes de la CMR en 3701 individuos (Estudios Pizarra, Segovia y Hortega), mediante la técnica SNPlex.

Fase 2: secuenciación de los genes (exones y secuencias intrónicas contiguas) asociados en la fase 1, por metodología Sanger, en 50 controles y 50 casos.

Fase 3: estudio de las variantes identificadas en la fase 2, en las muestras incluidas en la Fase 1, mediante secuenciación de nueva generación (NGS).

Fase 4: análisis de todos los polimorfismos previamente identificados con DMT2 y parámetros relacionados, en el estudio Di@bet.es, utilizando NGS.

Resultados: Se encontraron 17 asociaciones significativas de los polimorfismos analizados en las distintas poblaciones analizadas, y para los distintos parámetros relacionados con la DMT2. Cinco polimorfismos pertenecientes a estas asociaciones mostraron un comportamiento similar en la población de replicación, y se han validado sus efectos en el análisis de la población total de este trabajo. Las asociaciones confirmadas al analizar la población conjunta, han sido la asociación de SNPs del gen UQCRB y la glucemia en la población femenina, del gen NDUFS3 y

el riesgo de DMT2 en obesos; mientras que en los individuos no obesos, han sido las de los genes NDUFS2 y COX7C con la DMT2 y la del gen NDUFS5 con la PAD.

Conclusiones: Hemos encontrado asociaciones relevantes de los genes de la CMR con la DMT2 y parámetros relacionados, como la obesidad y la hipertensión. Estos resultados se han replicado en las diferentes muestras que incluye el presente estudio, lo que confirma la asociación de la CMR con el desarrollo de la DMT2. Se deberán realizar más estudios para verificar estas asociaciones y conocer los mecanismos por los que las diferentes variantes pueden actuar sobre el riesgo de desarrollar DMT2.

1. Introducción

1.1 Diabetes Mellitus

1.1.1 Descripción general y clasificaciones

La diabetes mellitus (DM) es un grupo de desórdenes metabólicos que afecta aproximadamente al 7-10% de la población mundial y cuya frecuencia está aumentando de forma rápida en todo el mundo (Imam 2012, Danaei y Finucane et al 2011). Según datos del International Diabetes Federation (IDF) del 2015, 415 millones de adultos sufren esta enfermedad (8.8% de la población mundial), aunque esta prevalencia no es homogénea (Figura 1). En España, el estudio Di@bet.es ha mostrado una prevalencia del 13.8 %, ajustada por edad y sexo (Soriguer et al 2012). La diabetes está caracterizada por la hiperglucemia crónica como manifestación clínica principal, aunque también se presenta en muchos casos asociada con alteraciones del metabolismo lipídico, proteico y de carbohidratos. La hiperglucemia, con el tiempo provoca daño orgánico a diferentes niveles, produciendo la disfunción y el fallo de diferentes órganos, especialmente ojos, riñón, nervios, corazón y vasos sanguíneos. En comparación con la población no diabética, los diabéticos poseen una expectativa de vida 7 años menor, un efecto relacionado directamente con las principales complicaciones de la enfermedad (Morgan et al 2000). Los pacientes con diabetes desarrollan complicaciones macro (enfermedad coronaria arterial y periférica vascular entre otras) y microvasculares (nefropatías, retinopatías y neuropatías periféricas) de forma mucho más frecuente que en la población general. En todo el mundo, la DM es la mayor causa de enfermedad coronaria arterial, amputaciones de miembros, enfermedad renal terminal y ceguera (Murea et al 2012).

Según las distintas etiologías que producen esta alteración, se clasifica a la DM de la siguiente manera (American Diabetes Association, 2012 y 2016):

- Diabetes Mellitus tipo 1 (DMT1): en el 90% de los casos se produce por procesos autoinmunes que llevan a la destrucción de las células beta del páncreas y finalmente a una deficiencia en la producción de insulina. Representan un 5-10% del total de los pacientes diabéticos.

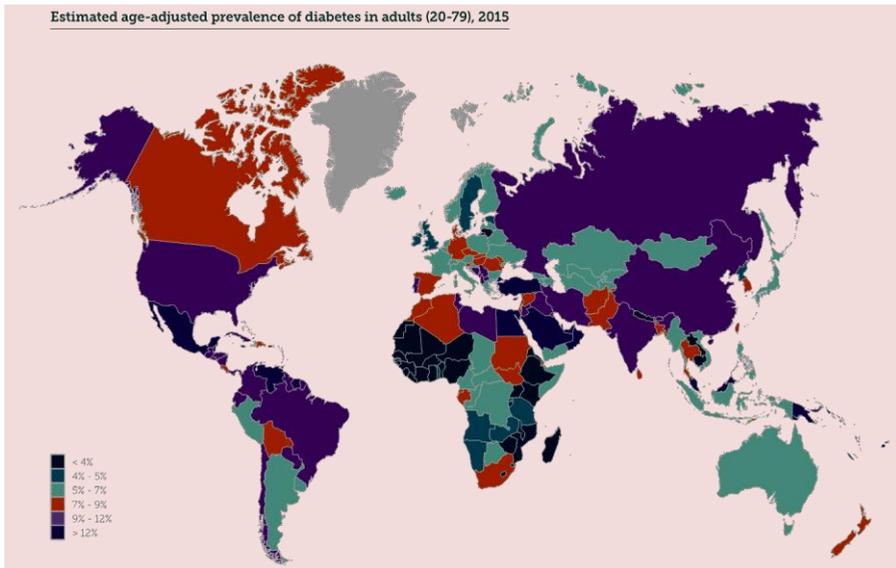


Figura 1. Prevalencia de Diabetes indicada por países. (IDF Diabetes Atlas, 7th edn. Brussels, Belgium: International Diabetes Federation, 2015).

- Diabetes mellitus tipo 2 (DMT2): es la forma más frecuente de esta enfermedad, representando alrededor del 80-90 % del total de las diabetes diagnosticadas. La insulino-resistencia (RI) es la principal característica de estos pacientes, presentándose en un 90% de los casos, conjuntamente con una pérdida progresiva de la secreción de la insulina.
- Diabetes mellitus gestacional (DG): diabetes desarrollada por algunas mujeres en el segundo o tercer trimestre de la gestación. La hiperglucemia se resuelve en la mayoría de las mujeres después del parto, si bien, haber padecido DG es un factor de riesgo independiente para el desarrollo de DMT2 en el futuro.
- Otros tipos específicos de Diabetes:
 - Endocrinopatías: Acromegalia, Síndrome de Cushing, Feocromocitoma, hipertiroidismo, Glucagonoma, entre otros.
 - Enfermedades del páncreas exocrino: pancreatitis, neoplasias, hemocromatosis, fibrosis quística, etc.
 - Debidas a alteraciones genéticas

-Defectos en la función de las células beta del páncreas: Ciertas formas de la diabetes están asociadas a defectos monogénicos en el genoma, que hacen que las células beta puedan sufrir una reducción en su metabolismo, en su regeneración, o en el número de estas células. La diabetes de la juventud de aparición en la madurez (MODY) es la más representativa y puede ser producida por alteraciones en varios genes. Ejemplos de éstas son las diabetes causadas por mutaciones en el gen HNF4 α (MODY1), GCK (MODY2), HNF1 α (MODY3), PDX1 (MODY4), HNF1 β (MODY5), NEUROD1 (MODY6), siendo las tres primeras las más frecuentes. Se han descrito más de 600 mutaciones diferentes a lo largo del gen GCK en el mundo y más de 200 en la región codificante y promotora en el gen HNF1 α . Las variaciones genéticas en el gen HNF4 α son más raras, hallándose hasta el momento sólo unas 31 (Osbak et al 2009, Ellard y Colclough 2006, Vaxillaire et al 2012). Los defectos genéticos en los genes INS, ABCC8 y KCNJ11 se encuentran en menos del 1% de los MODY, y las de IPF1 y CEL son muy poco frecuentes. En el ADN mitocondrial también se han visto alteraciones afectando la función de las células beta, provocando la llamada diabetes mitocondrial. Mutaciones en distintos genes que codifican para proteínas nucleares, tales como factores de transcripción, pueden causar defectos en el desarrollo y en la función pancreática, provocando como consecuencia una variedad de fenotipos que van desde la agenesia pancreática con diabetes neonatal e insuficiencia exocrina, hasta la diabetes de aparición en la adolescencia o en edad adulta. Otros niveles a los que se han encontrado defectos monogénicos causantes de distintos tipos de diabetes son en la incorporación de glucosa en el organismo, el metabolismo celular, a nivel lisosomal y del retículo endoplásmico, síntesis y secreción de insulina o diabetes autoinmune monogénica (Schwitzgebel 2014). En la figura 2 se muestran a qué niveles pueden encontrarse estas mutaciones en las células beta.

-Defectos en la acción de la insulina: resistencia a la insulina tipo A, síndrome de lipodistrofia, síndrome de Rabson-Mendenhall, Síndrome de Donohue (Leprechaunismo).

-Síndromes genéticos asociados con diabetes: Síndromes de Down, de Turner, de Klinefelter, de Wolfram, de Prader Willi, Porfiria, entre otros.

- Diabetes inducida por fármacos: glucocorticoides, hormona tiroidea, diazoxida, agonistas β -adrenérgicos, pentamidina, etc.
- Infecciones: Cytomegalovirus, Rubeola congénita, etc.
- La diabetes latente autoinmune del adulto (LADA) es un tipo de diabetes que ha generado controversia para ubicarla en la clasificación general de la American Diabetes Association (ADA). Este tipo de diabetes tiene un componente autoinmune, ya que los pacientes presentan anticuerpos contra las células de los islotes pancreáticos, pero a su vez tienen un fenotipo similar al de la Diabetes mellitus tipo 2 (Naik et al 2009, Pollak et al 2012). Su aparición es en edades superiores a 35 años y responden bien al tratamiento sin necesidad de insulina en la primera etapa de la enfermedad. Se le ha dado el nombre de Diabetes mellitus tipo 1 de lenta progresión (Kobayashi et al 1993) o latente (Groop et al 1986) o también diabetes tipo 1.5 (Palmer y Hirsch 2003).

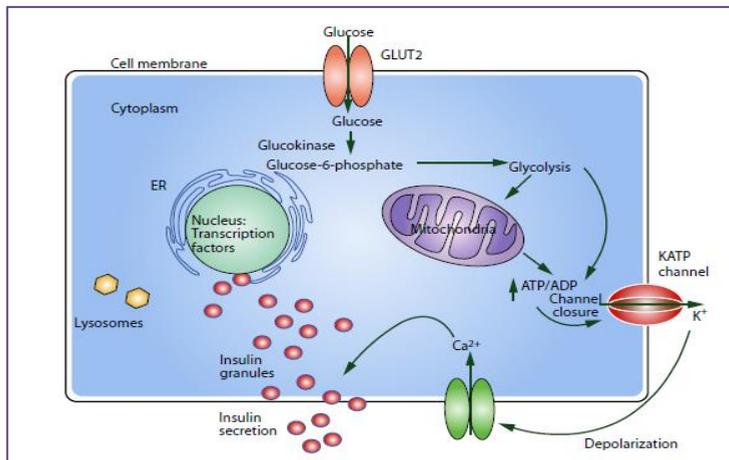


Figura 2. Esquema de una célula beta. Localización de los pasos en los que se han identificado mutaciones que producen diabetes monogénicas (J Diabetes Investig. 2014; 5(2):121-33).

En la figura 3 se representan los porcentajes de los tipos más frecuentes de diabetes. La diabetes mellitus también puede clasificarse en primaria, cuando la causa es genética o inmunológica y el daño es directamente en la célula beta pancreática, o en diabetes secundaria, si es producida por otras enfermedades que atacan al páncreas.

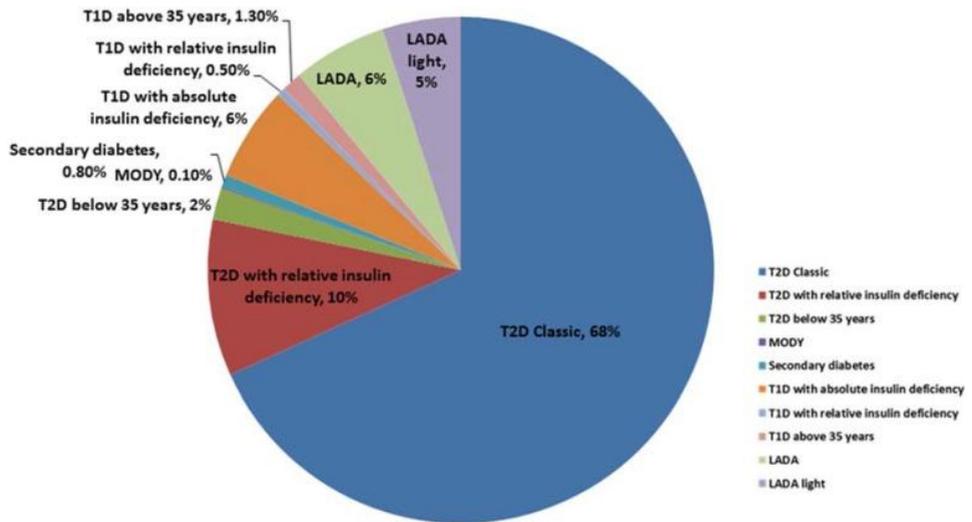


Figura 3. Proporciones de las distintos tipos de diabetes. El gráfico está elaborado a partir de los datos del ANDIS project (All New Diabetics in Scania, <http://andis.ludc.med.lu.se>) de abril de 2012.

Desde un punto de vista genético, son tipos de diabetes primarias las diabetes monogénicas, como la diabetes mitocondrial o las MODY; las diabetes multifactoriales, que comprenden mayoritariamente a la DMT1 y 2; y las ocasionadas por síndromes genéticos como los Síndromes de Down, Turner, Klinefelter, Wolfram y Prader Willi (Murphy et al 2008). Las diabetes provocadas por endocrinopatías o ciertos fármacos son ejemplos de diabetes secundarias (figura 4).

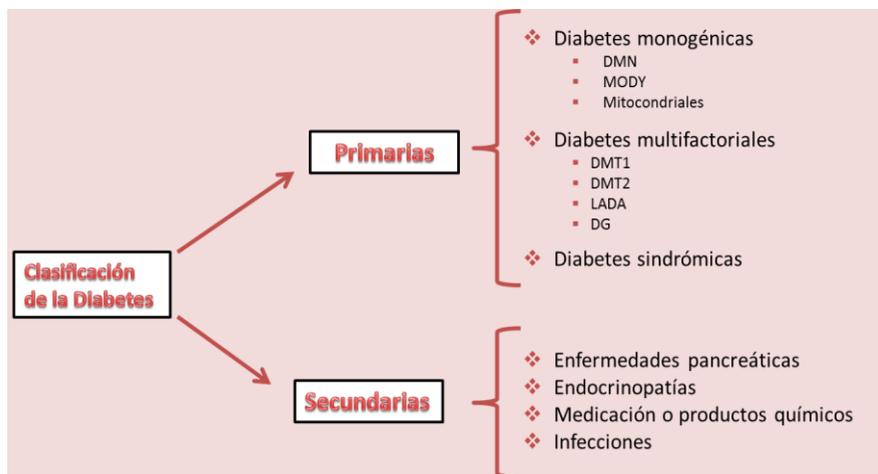


Figura 4. Clasificación de la diabetes en diabetes primaria y secundaria. DMN: Diabetes mellitus neonatal.

1.1.2 Criterios diagnósticos de la diabetes

Según las recomendaciones de la Organización Mundial de la Salud (OMS) y la IDF del 2006, los criterios diagnósticos para la diabetes e hiperglucemias intermedias están basados en los valores de glucemia en ayunas y en la prueba de tolerancia a la glucosa. En el 2009, la IDF, la European association for the study of Diabetes (EASD) y representantes de la ADA recomendaron el uso de la hemoglobina glicosilada A1C (HbA1C) para diagnosticar la diabetes, con un valor de corte del 6.5%. La determinación de la HbA1C debe hacerse a través de un método certificado y estandarizado, por ejemplo, por el National Glycohemoglobin Standardization Program (NGSP) y por el Diabetes Control and Complications Trial (DCCT). En estos momentos este tipo de ensayos no están ampliamente extendidos, si no que se usan de rutina otros métodos no suficientemente fiables para propósitos diagnósticos, más aún en ciertas regiones geográficas (American Diabetes Association 2011 y 2014).

Los criterios mundialmente utilizados para diagnosticar la diabetes mellitus, diferenciándola de otros tipos de alteraciones de la glucemia, son los indicados en la tabla 1.

| Diagnóstico | Glucemia en ayunas | Glucemia a las 2 horas de la PTG * | HbA1C |
|----------------------------------------|------------------------------------|------------------------------------------------|----------|
| Diabetes | ≥ 7.0 mmol/l (126 mg/dl) | ≥ 11.1 mmol/l (200 mg/dl) | ≥ 6.5%** |
| Tolerancia alterada a la glucosa (TAG) | < 7.0 mmol/l (126 mg/dl) | ≥ 7.8 mmol/l y < 11.1 mmol/l (140 a 200 mg/dl) | |
| Glucosa alterada en ayunas (GAA) | 6.1 a 6.9 mmol/l (110 a 125 mg/dl) | < 7.8 mmol/l (140 mg/dl) | |

Tabla 1. Criterios diagnósticos de alteraciones del metabolismo glucídico. PTG: prueba de tolerancia a la glucosa, ingesta de una carga oral de glucosa de 75 g, y medición de la glucemia en ayunas a las 2 horas. La presencia de cualquiera de los 3 criterios nombrados en la tabla es suficiente para diagnosticar Diabetes. *Si la glucemia a las 2 horas de la PTG no se mide, no se puede excluir el diagnóstico de tolerancia alterada a la glucosa (TAG) y el diagnóstico de Diabetes no puede confirmarse. ** Esta determinación debe realizarse usando un método certificado y estandarizado por el NGSP o por el DCCT para utilizarse en el diagnóstico.

La ADA recientemente ha agregado como otro posible criterio diagnóstico de diabetes mellitus, a la glucosa en plasma medida en forma aleatoria mayor a 11.1 mmol/l (200 mg/dl) en pacientes con síntomas clásicos o crisis de hiperglucémica (ADA 2015).

Cuando el diagnóstico de la diabetes está establecido, para poder clasificar el tipo de diabetes que tiene el paciente, es necesario revisar la historia clínica y determinar otros parámetros bioquímicos. Distintas características se deben tener en cuenta a la hora de hacer la clasificación de las diabetes más comunes y de los MODY, por ejemplo, como la edad de aparición, presencia de obesidad, entre otras (Tallapragada et al 2015).

1.2 Diabetes Mellitus tipo 2

1.2.1 Prevalencia de la diabetes mellitus tipo 2

La diabetes mellitus tipo 2 (DMT2), antes también llamada diabetes no insulino-dependiente, representa aproximadamente el 80-90% de los pacientes diabéticos en todo el mundo. El crecimiento en la población mundial de la incidencia de esta enfermedad, se debe en gran medida, al aumento en el consumo de dietas ricas en grasas y en los bajos niveles de actividad física, lo que contribuye a una mayor obesidad (Zimmet et al. 2001).

Más de 330 millones de personas en todo el mundo padecen DMT2, 60 millones de ellas pertenecientes a Europa (<http://www.who.int/diabetes>). En la mayoría de los países europeos se ha visto una prevalencia moderada, con menos de un 10 % en la población de menos de 60 años y un 10-20 % en personas entre los 60-80 años (DECODE-Study-Group 2003). En España, el Spanish Insulin Resistance Study (SIRS) fue uno de los primeros en incluir participantes de varias ciudades españolas (Arévalo, Talavera de la Reina, Guadalajara, La Coruña, Avilés, Vic, Alicante y Mérida), hallando una prevalencia de DMT2 del 10,2% (según criterios de la organización mundial de la salud,OMS, 1999) en el 2001. Estudios regionales han estimado sus prevalencias, encontrándose unos valores entre 10 a 15% en el 2007 (Valdés et al 2007). Además, se ha realizado un estudio comparativo de las

prevalencias de algunas comunidades autónomas recopiladas en las encuestas nacionales de salud (ENS) de 2001 y 2003 (Basterra-Gortari y Martínez-González 2007). Canarias es la comunidad con mayor prevalencia de diabetes, seguida de Andalucía y Baleares. Tomando Madrid como referencia, Canarias y Andalucía han mostrado una prevalencia de diabetes significativamente superior a la de Madrid. Por otro lado, La Rioja, Asturias, Navarra, Castilla y León y País Vasco tienen una prevalencia significativamente inferior. La encuesta europea de salud en España (2009 y 2014), confirma la existencia de esta distribución en las comunidades autónomas (https://www.msssi.gob.es/estadEstudios/estadisticas/EncuestaEuropea/Enc_Eur_Salud_en_Esp_2014_datos.htm, Estrategia en Diabetes del Sistema Nacional de Salud 2012). El estudio Di@bet.es es el primer estudio a nivel nacional en España que analiza la prevalencia de la diabetes y de alteraciones del metabolismo de la glucosa en adultos (Soriguer et al 2012). En este estudio se ha visto que casi un 30 % de los participantes tienen alguna alteración en el metabolismo de la glucosa, con una prevalencia de DM, ajustada por edad y sexo, del 13.8 % en adultos. Aproximadamente la mitad de estos pacientes no sabían que tenían la enfermedad. A su vez, el 9.2 % de pacientes presentan TAG y un 3,4 % GAA. Este estudio ha confirmado que la prevalencia de la DM aumentaba significativamente con la edad ($p < 0.0001$), y que es mayor en hombres que en mujeres ($p < 0.001$). En el 2014, se realizó un estudio de la prevalencia de la diabetes y de varias enfermedades relacionadas, al comparar dicho parámetro entre Andalucía y el resto de comunidades autónomas, analizando los pacientes del estudio Di@bet.es (Valdés et al 2014). La prevalencia de la diabetes es del 16.3% en Andalucía, mientras que en el resto de España es del 12.5%.

1.2.2 Fisiopatología de la DMT2

La fisiopatología de la DMT2 se centra en la insulina y los mecanismos por los cuales los niveles de glucosa en sangre (glucemia) son controlados. Las células beta pancreáticas sintetizan insulina sin tener en cuenta los niveles de glucosa sanguínea, la almacena en vacuolas y es liberada frente a una elevación de la glucemia. La insulina es el principal regulador de la incorporación de glucosa por las células y,

por tanto, de sus niveles en sangre. La insulina induce la conversión de glucosa en glucógeno, para su almacenamiento interno en el hígado y en las células musculares esqueléticas. Además, provoca la inhibición de la producción de glucosa en el hígado, y a nivel del tejido adiposo, actúa inhibiendo la lipólisis en los adipocitos (Figura 6).

Una caída de la glucemia provoca una disminución en la liberación de insulina y un aumento de la liberación de glucagón a partir de las células alfa pancreáticas, estimulándose así la conversión de glucógeno en glucosa (Lin y Sun 2010).

Un mecanismo importante para el desarrollo de DMT2 es la resistencia a la insulina (RI). Esta se produce cuando las células presentan una respuesta reducida frente a la acción de esta hormona, o hay ausencia de dicha respuesta. Por este motivo, la célula beta debe aumentar la producción y los niveles en sangre de insulina para compensar la RI, y controlar así los niveles de glucosa. La hiperglucemia, el exceso de producción de insulina y otros factores provocan estrés a nivel de las células beta.

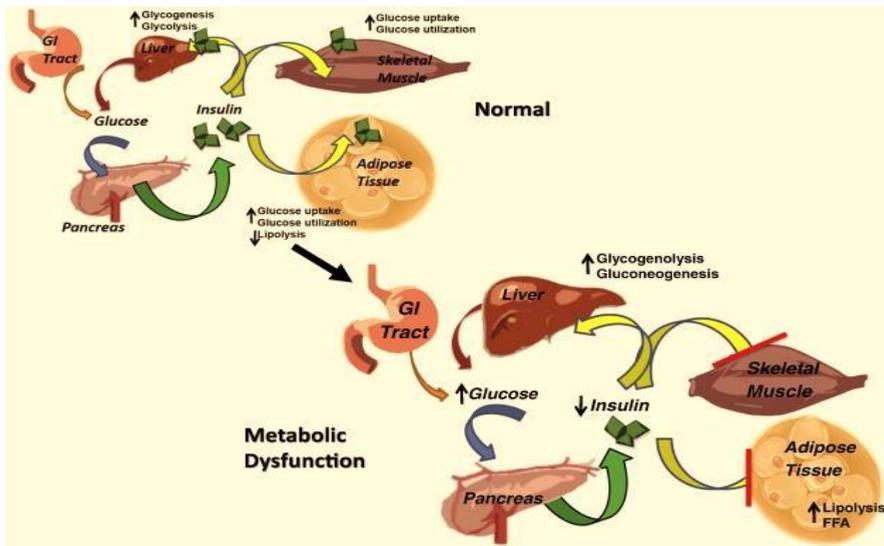


Figura 6. Patofisiología de la hiperglucemia y del aumento de los ácidos grasos en circulación en la DMT2. (Biochim Biophys Acta. 2016, pii: S0925-4439(16)30187-9).

Dicho estrés puede alterar la producción y la liberación de insulina e, incluso, provocar la muerte de estas células. Si se produce un exceso de muerte celular frente

a la tasa de regeneración, el páncreas dejará de producir suficiente insulina con el tiempo; lo que termina provocando hiperglucemia y DMT2 (Ramalingam et al 2016). Los tres pasos claves en la aparición de la DMT2 son los siguientes (Stumvoll et al 2005):

- Acción alterada de la insulina
- Producción hepática aumentada de la glucosa
- Disminución de la secreción de insulina

1.2.3 Factores que afectan el desarrollo de la DMT2

La DMT2 es un desorden metabólico que se produce al combinarse factores de riesgo ambientales con una susceptibilidad genética de padecer dicha enfermedad (figura 7). Dentro de los factores ambientales, podemos encontrar factores no modificables, como la edad, grupo étnico, antecedentes familiares de diabetes, y factores modificables, como lo son la obesidad o el sedentarismo. La obesidad es sin duda el factor de riesgo más importante para el desarrollo de DMT2, seguido por la inactividad física (Perry et al 2002). Los factores genéticos pueden influir directamente en la aparición de la DMT2 o pueden afectar a otros factores de riesgo como por ejemplo, la obesidad.

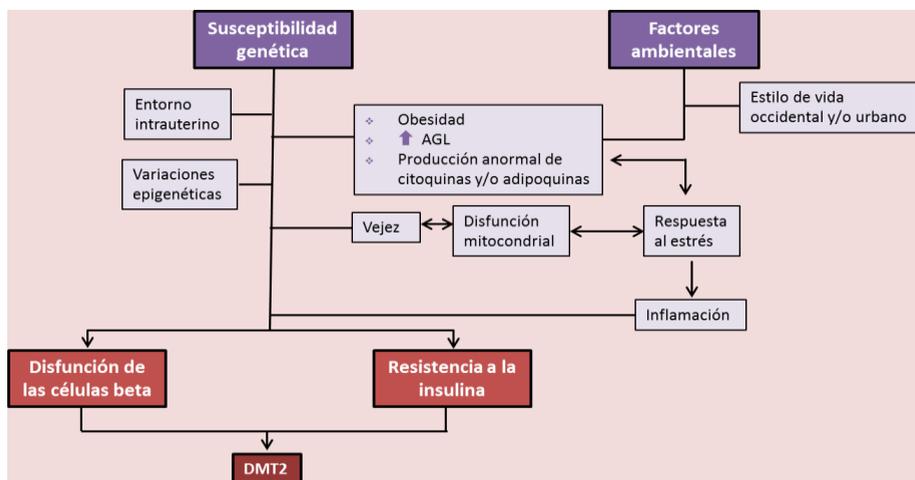


Figura 7. Factores de riesgo de la DMT2. Relaciones entre los distintos factores que influyen en la aparición de la DMT2. AGL: ácidos grasos libres. (Basado en el artículo de 2014, World J Diabetes. 2014; 5(2):128-40).

1.2.3.1 Factores genéticos de la DMT2

Se ha estimado que entre el 30% y el 70% del riesgo de padecer DMT2 puede ser atribuido a factores genéticos (Poulsen et al 1999). Los parientes en primer grado de pacientes con DMT2 tienen 3 veces más riesgo de desarrollar esta enfermedad que individuos sin historia familiar de DMT2 (Florez et al 2003). Este aumento del riesgo, hallado en varios estudios familiares, puede no ser exclusivamente debido a factores genéticos que supongan cambios en la secuencia del ADN (Ali 2013). Parecen existir también procesos epigenéticos heredables y factores relacionados con el embarazo y el entorno intrauterino que pueden modificar el riesgo de DMT2 (Slomko et al 2012).

La naturaleza multifactorial de la DMT2 ha hecho que la identificación de variaciones genéticas relacionadas con ella sea un reto complejo. Los principales métodos que pueden permitir identificar loci de susceptibilidad para una enfermedad o rasgos fenotípicos, han evolucionado rápidamente en los últimos años. En un principio, los análisis de ligamiento en familias y los estudios de asociación de genes candidatos, han sido utilizados y se están utilizando aún para este propósito, estableciéndose con ellos la importancia de los factores genéticos en la etiología de la DMT2 (Imamura y Maeda 2011). En los últimos años, se han realizado estudios de asociación de genoma completo (GWAS) que abarcan el estudio de marcadores genéticos a lo largo de todo el genoma. Inicialmente se realizaron en pequeñas muestras y, actualmente, se están realizando con tamaños muestrales grandes.

La secuenciación de nueva generación (NGS) ha permitido el estudio del exoma y del genoma, analizándose así de forma directa la secuencia genética.

Se han realizado, y se siguen haciendo, grandes esfuerzos para dilucidar la genética de la DMT2 mediante las diferentes estrategias disponibles en cada momento, aunque hace solo 10 años que se han obtenido progresos sustanciales en la identificación de las variantes genéticas asociadas fuertemente con esta enfermedad (Kato et al 2013). Pese a ello, aún no se ha identificado la mayoría del componente genético de la DMT2, estimándose que falta por identificar el 85-90%.

Estudios de ligamiento

Los estudios de ligamiento son útiles para identificar loci de alta penetrancia y con un modo de herencia conocido, pero no así para identificar genes causales de enfermedades complejas como la DMT2. Sólo dos genes han sido identificados por ligamiento a pesar de todos los esfuerzos realizados en estudios sobre la DMT2, el gen de la calpaina 10 (CAPN10) y el del factor de transcripción 7 like, 2 (T-cell specific, HMG-box, TCF7L2). El CAPN10 ha sido el primero en asociarse a DMT2 en esta clase de estudios, identificándose primero la región ligada a DMT2 (Hanis et al 1996) e confirmando posteriormente a este gen como el responsable de la asociación (Horikawa et al 2000). En estos momentos, la función de este gen en el metabolismo de la glucosa y en la patogenia de la DMT2 es desconocida, a pesar de que varios trabajos han confirmado su asociación (Song et al 2004). TCF7L2 ha sido identificado como gen de susceptibilidad de DMT2 en poblaciones danesas, estadounidenses e islandesa, a partir de la fuerte asociación de obtenida en estudios de ligamiento del cromosoma 10q en una población mejico-americana (Duggirala et al 1999). Varios SNPs en este gen se han asociado con DMT2 en estudios GWAS en diferentes grupos étnicos. La asociación de TCF7L2 continúa siendo en estos momentos la más replicada a nivel mundial y es la asociación más fuerte encontrada hasta el momento (Tong et al 2009).

Estudios de genes candidatos

En este tipo de estudios, se seleccionan los genes que por su función, pueden jugar un papel en la patogénesis de la DMT2, mediante la presencia en ellos de variantes genéticas. En general, se eligen genes involucrados en el metabolismo de la glucosa, en la secreción de insulina, en el metabolismo lipídico, en la señalización de la insulina, genes relacionados con las terapias antidiabéticas, etc, todos ellos sistemas relacionados con la DMT2. Los estudios de genes candidatos representan una de las mejores herramientas para hallar asociaciones con enfermedades, dado que las asociaciones encontradas pueden tener una base patogénica y funcional directa (Brunetti et al 2014). Mediante esta aproximación se han identificado

principalmente 8 genes de gran relevancia: el gen que codifica para el receptor gamma activado por el proliferador del peroxisoma (PPARG), el gen del sustrato del receptor de la insulina 1 y 2 (IRS1 y 2), el miembro 11 de la subfamilia J de los canales potasio dependientes (KCNJ11), el gen de la wolframina (WFS), y los factores nucleares hepáticos 1 homeobox A (HNF1A), B (HNF1B) y el 4 alfa (HNF4A). Algunos de estos genes codifican proteínas que son la diana de fármacos antidiabéticos: tiazolidinedionas y sulfonilreas actúan sobre PPARG y KCNJ11, respectivamente (Ahlqvist et al 2011). KCNJ11 juega un papel importante en la regulación de la secreción de insulina en las células beta (Hani et al 1998). Mutaciones activadoras de su función son causales de diabetes neonatal. Polimorfismos del gen PPARG, especialmente Pro12Leu, provocan un aumento en el riesgo de padecer diabetes de un 20% (Ruchat et al 2009). El síndrome de Wolfram 1 se caracteriza por diabetes insípida y juvenil, entre otros fenotipos; mutaciones en el gen WFS-1 lo causan, y ese gen además parece estar involucrado en la función de las células beta (Minton et al 2002). Los genes HNF1A, HNF1B y HNF4A son genes causales de MODY y juegan un papel importante en la regulación de las funciones metabólicas hepáticas, en el desarrollo del hígado y de las células beta del páncreas. Se han encontrado variantes de estos genes que se asocian con la disminución de la secreción de insulina y con el aumento del riesgo de padecer DMT2 en varias poblaciones, pero que no son causales de MODY (Hegele et al 2003). En la figura 8 se enumeran distintos genes relacionados con el metabolismo de la glucosa, asociados con DMT2.

Estudios de asociación de genoma completo (GWAS)

Un avance importante en el entendimiento de las bases genéticas de la DMT2 se produjo con la llegada de los GWAS (Voight et al 2011). Estos estudios se basaron originariamente en la hipótesis de que el riesgo genético para enfermedades complejas comunes es debido, normalmente, a variaciones comunes, es decir, que tienen una frecuencia alélica mayor del 5 %. Con esta metodología, estudiando en la mayoría de los casos, variantes genéticas comunes a lo largo de todo el genoma, se han identificado unos 70 loci de susceptibilidad para la DMT2 en poblaciones

europas y asiáticas (Kato et al 2013). También se han descubierto más de 40 de estas variantes asociadas con distintos fenotipos o parámetros asociados con la DMT2. Estos estudios han permitido identificar genes y rutas involucrados en la patogenia de la DMT2 de las que se desconocía su posible implicación en esta enfermedad (Pranavchand y Reddy 2016).

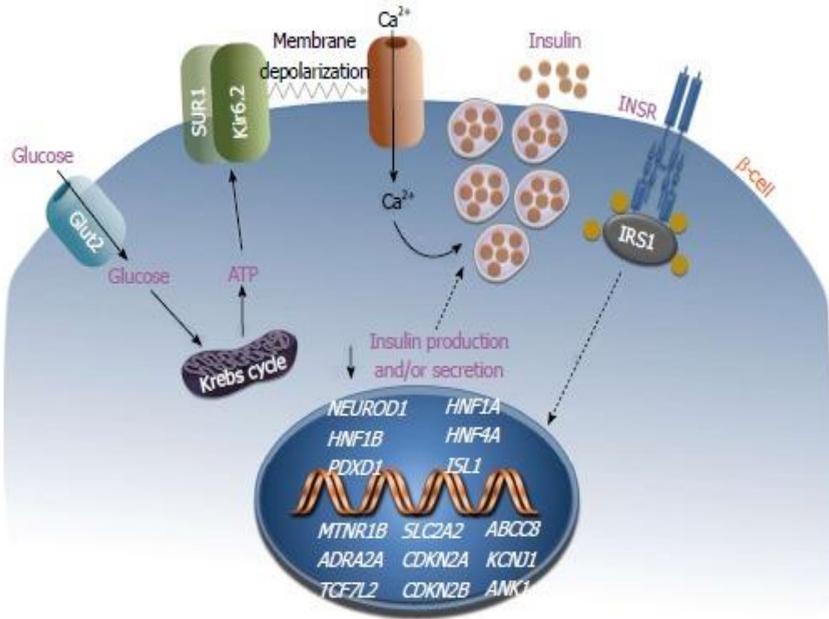


Figura 8. Representación esquemática de la célula beta del páncreas. Se indica con flechas el mecanismo por el cual se secreta la insulina ante un exceso de entrada de glucosa a la célula. Este se ve alterado por la presencia de variantes relacionadas con DMT2, reduciéndose la secreción de insulina. En blanco se representan los genes asociados con DMT2 (World J Diabetes. 2014; 5(2):128-40).

Muchos de los loci hallados confirmaron los resultados que se habían encontrado en los estudios de genes candidatos. Una gran cantidad de los polimorfismos asociados con DMT2 están ubicados en regiones intrónicas, intergénicas, promotoras, u otras zonas genómicas que no pertenecen al exoma. Otros GWAS han identificado loci asociados con parámetros importantes para la aparición y el diagnóstico de la

DMT2, como son la glucemia y la insulinemia en ayunas, la hemoglobina glicosilada A1C (HbA1C) sanguínea, la RI y la TAG (Basile et al 2014). De esta manera, se han identificado loci asociados con diferentes fenotipos clave en la etiología de la DMT2: los relacionados con la secreción de insulina, con la RI o con la obesidad, entre otros (figura 9).

En el European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) han confeccionado un resumen detallado en su página web, de los distintos estudios GWAS de DMT2 recopilados hasta el momento (<http://www.ebi.ac.uk/gwas/search?query=type%20%20diabetes>).

A pesar de que se han hecho muchos hallazgos, los loci asociados en GWAS representan una pequeña proporción de la heredabilidad estimada, y la mayoría tienen un efecto muy modesto sobre el riesgo de desarrollar la enfermedad, que va de un 10 a un 15%.

Estos estudios pese al gran adelanto que han supuesto, presentan importantes limitaciones como son la falta de poder estadístico para identificar variantes raras, la gran cantidad de marcadores cuya asociación es dudosa, y su elevado coste. Por otro lado, muchos marcadores no han sido replicados en otros estudios y los riesgos relativos encontrados son muy bajos.

Estudio de variantes raras

La falta de explicación de la mayor parte de las bases genéticas de la DMT2 (Manolio et al 2009, Ali 2013), apoyan la idea de que las variantes raras (MAF < 1%) de alta penetrancia podrían estar contribuyendo al aumento del riesgo de padecer enfermedades comunes como la DMT2 (Pritchard y Cox 2002).

Algunas plataformas para GWAS, recientemente han estado incorporando dentro de sus marcadores, a variaciones con MAFs más bajos, que llegan al 1%. Las más ampliamente utilizadas son las de Illumina para exoma o genoma completo (www.Illumina.com/techniques/microarrays/human-genotyping.html).

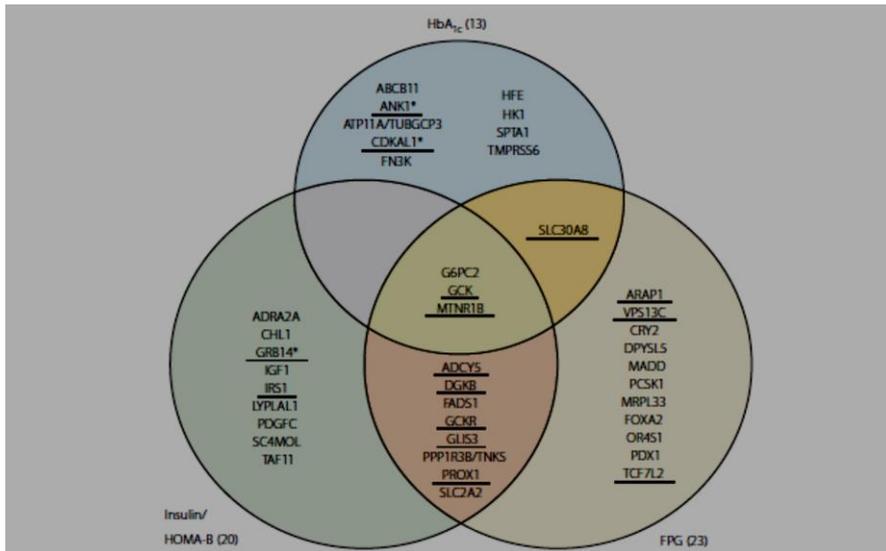


Figura 9. Esquema de los genes asociados en estudios GWAS, que se han relacionado con otros parámetros de la DMT2. Los genes subrayados en negro se han asociado además con DMT2. FTG: glucosa plasmática en ayunas; HbA_{1c}: Hemoglobina glicosilada; HOMA-B: modelo homeostático de la función de las células beta (J Diabetes Investig. 2013; 4(3):233-44).

Estos estudios pretenden identificar variantes de baja frecuencia con un gran efecto sobre el riesgo de desarrollar DMT2, con el fin de facilitar la identificación de los mecanismos por los que modularía el desarrollo de la DMT2 (Steinthorsdottir et al 2014, Cirulli y Goldstein 2010).

Se han identificado variantes raras de los genes CCDN2, PAM and PDX1 asociadas con la DMT2 (Steinthorsdottir et al 2014). Otro hallazgo interesante son las 12 mutaciones con pérdida de función del gen SLC30A8, que contribuirían a disminuir el riesgo de DMT2 (Flannick et al 2014).

La secuenciación del exoma de PPARG de casos y controles de varios estudios poblacionales y de varios grupos étnicos diferentes han permitido identificar variantes raras asociadas a una disminución de la actividad de la diferenciación del adipocito y un riesgo de DMT2 aumentado (Majithia et al 2014).

Si bien estas tecnologías pueden permitir identificar cambios genéticos que contribuyan a la DMT2, sigue siendo un gran reto a nivel económico, y obtener un

tamaño muestral adecuado para la detección de variantes de baja frecuencia es complicado.

Estudios de interacciones entre genes y factores ambientales

Además de lo ya indicado, hay que considerar más factores que pueden modular el riesgo de DMT2 como son las interacciones gen-gen y entre el ambiente y las diferentes variantes. Su análisis podría ayudar a identificar nuevas rutas con diferentes efectos sobre el riesgo de padecer DMT2. La identificación de loci asociados con esta enfermedad han dado pie a la realización de muchos estudios para buscar interacciones entre esos loci y los factores de riesgo ambientales para la DMT2 (Franks et al 2013 y Cornelis y Hu 2012). Una de las primeras publicaciones de este tipo se centró en el gen de la dioxigenasa dependiente del alfa cetoglutarato (FTO), el SNP rs9939609, relacionándolo con la actividad física (Andreasen et al y Rampersaud et al 2008). En este estudio se observó que el efecto de esta variación de FTO sobre el índice de masa corporal (IMC) y sobre la acumulación de tejido adiposo es modificado por la actividad física. Sin embargo, otros trabajos han puesto en duda estos resultados, debido al hallazgo de una influencia étnica y/o de la dieta en ellos. Un primer estudio de la interacción entre los loci asociados a la DMT2 y la actividad física incluyó 17 loci, de los cuales uno, perteneciente al gen HNF1B, interactuaba con el nivel basal de actividad física (Brito et al 2009). Se ha encontrado un efecto de la ingesta de grasas y carbohidratos de la dieta sobre la asociación entre la variante rs10423928 del gen del receptor del polipéptido inhibitorio gástrico (GIPR) y la DMT2 (Sonestedt et al 2012). Más trabajos han indagado en la interacción entre genes relacionados con la DMT2 y el ambiente (Cornelis et al 2009, Nettleton et al 2010 y Qi et al 2009) pero todavía quedan por analizar muchas posibles interacciones entre factores genéticos y ambientales. Varios estudios han demostrado que el efecto de algunas variantes genéticas sobre el desarrollo de la DMT2 difiere en base a la presencia o no de obesidad, (Perry et al. 2012 y Morris et al 2012, Wu et al 2014) o en función del género (Alkhalaf et al 2015).

Epigenética y DMT2

Grandes avances se han hecho en el estudio de la relación de la epigenética con la patología de la DMT2. Muchos factores ambientales que influyen de manera decisiva en la aparición de esta enfermedad son capaces de inducir también cambios epigenéticos, que, al alterar la función de ciertos genes, podría llevar a la aparición de la DMT2. Por ejemplo, se sabe que el riesgo de desarrollar DMT2 va aumentando con la edad, y de la misma manera lo hacen los errores en la metilación del ADN, por lo que éstos podrían contribuir a la DMT2 (Calvanese et al 2009). Varios estudios han analizado el estado de metilación de genes relacionados con la DMT2, y con otro factor ambiental muy importante, la obesidad. Se han hallado variaciones en el grado de metilación en esos genes, asociándolos con estas enfermedades (Raciti et al 2015).

Cambios en el metabolismo de los grupos metilo, como el aumento de los niveles de homocisteína, el consumo de dietas pobres en azufre que provocan aumento de metionina y mayor síntesis de cisteína, y los polimorfismos en la metileno-tetrahidrofolato reductasa, predisponen para el desarrollo de DMT2 (Wren y Garner 2005).

Los miARNs son esenciales para el normal desarrollo del páncreas y están implicados también en la diabetes y en la regulación de la acción de la insulina (Poy et al 2004, y Rottiers y Näär 2012). Una reciente revisión detalla los estudios en los cuales la expresión de diferentes miARNs está alterada en distintos tejidos durante el desarrollo de la diabetes (Guay et al 2011). Otros estudios muestran que ciertos miARN estaban en niveles elevados en suero de pacientes con DMT2 y parecían estar involucrados en la función celular y en la DMT2 (Kong et al 2011). Con la inflamación, la hiperglucemia y la lipotoxicidad como factores clave en la patología de la DMT2, y como factores predisponentes de la disfunción de la célula beta, los miARN podrían explicar parte de la relación entre el metabolismo graso alterado, la hiperglucemia crónica, la inflamación (aumento de citoquinas pro-inflamatorias) y la apoptosis de la célula beta (Gilbert y Liu 2012). Un ejemplo es el aumento del

miARN codificado por el citomegalovirus humano aislado de plasma, HCMV-miR-UL112 (Li 2011) que aumenta el riesgo de desarrollar DMT2.

1.2.3.2 Factores de riesgo de la DMT2

A continuación se enumeran los factores de riesgo ambientales y adquiridos que aumentan el riesgo de padecer DMT2:

- Grupo étnico
- Edad superior a 45 años
- Antecedentes familiares de diabetes
- Diabetes durante un embarazo previo
- Madre de recién nacido de más de 4 kg
- Peso corporal excesivo (obesidad y obesidad central)
- Dieta alta en grasas
- Niveles sanguíneos de triglicéridos >250 mg/dl y colesterol HDL < 35 mg/dl
- Hipertensión arterial (PA superior o igual a 140/90 mmHg)
- Trastorno en la tolerancia a la glucosa
- Sedentarismo
- Síndrome metabólico
- Poliquistosis ovárica
- Acantosis pigmentaria

Ciertos grupos étnicos, como los afroamericanos, hispanos, asiáticos, nativos norteamericanos y de las islas del pacífico sur, tienen un mayor riesgo de padecer DMT2, siendo éstos dos últimos los de mayor riesgo entre todos (Wilson 2007, <http://www.who.int/genomics/about/Diabetis-fin.pdf>).

En general, se sabe que con el aumento de la edad aumenta el riesgo de padecer DMT2. En personas mayores de 45 años este riesgo es considerable, por lo que debe controlarse el nivel de glucosa en la sangre al menos cada tres años a partir de esa edad.

El aumento de adiposidad es el principal factor de riesgo modificable de la DMT2, correlacionándose el aumento del IMC y de la obesidad con el riesgo de DMT2 (Cornelis y Hu 2012). Tanto los individuos de peso adecuado como los obesos, se

benefician del efecto protector de la actividad física en la DMT2, lo que sugiere que este otro factor es independiente del grado de sobrepeso. Una ingesta calórica excesiva y una dieta de baja calidad nutricional son otros factores independientes importantes en el aumento de la DMT2.

La historia familiar de diabetes se ha asociado con una gran cantidad de desórdenes metabólicos asociados a la DMT2 y es un factor de riesgo para el desarrollo de la DMT2 (InterAct Consortium. 2013). Este mayor riesgo de padecer DMT2 podría estar mediado en parte por los componentes genéticos y ambientales que comparten los familiares, pero los factores precisos que intervienen en el aumento de este riesgo no se conocen exactamente. El riesgo de desarrollar DMT2 a lo largo de la vida es de un 40% para individuos que tienen un padre con este tipo de diabetes, mientras que sube a un 70 % si los dos padres son diabéticos (Tillil y Köbberling 1987).

Un factor, actualmente en discusión, es el papel que puede tener la contaminación (Murea et al 2012) en la aparición de la diabetes. Existen datos epidemiológicos de que una exposición crónica a contaminantes orgánicos del suelo como los pesticidas y herbicidas, altera al metabolismo glucídico e induce resistencia a la insulina. Una relación dosis-dependiente entre las concentraciones séricas de pesticidas orgánicos y la prevalencia de diabetes ha sido demostrada con odd ratios (OR) de 11.5 (Lee et al 2006). Por otro lado, la polución aérea podría estar involucrada en la etiología de la DMT2 a través de mecanismos que incluyen el estrés oxidativo, inflamación, disfunción endotelial, disfunción mitocondrial, entre otros (Rajagopalan and Brook, 2012 y Liu et al 2013), provocando un mal funcionamiento de la señalización de la insulina (Xu et al 2013). El estudio “Air Pollution and Lung and Heart Diseases in Adults, SAPALDIA” ha mostrado evidencias de que la exposición de larga duración a contaminantes ambientales como PM10 (materia particulada con diámetro menor a 10 μm) y el NO₂ (dióxido de nitrógeno) se asocian con DM a concentraciones por debajo de las guías de calidad del aire (Eze et al 2014). Además, se han encontrado relaciones entre metales como el arsénico, mercurio y el cadmio con la diabetes (Kuo et al 2013), y otros elementos no metales como el selenio (Galan-Chilet et al, 2017).

1.2.3.2.1 Obesidad

La obesidad (OB) es una condición metabólica caracterizada por un exceso de acumulación lipídica. La OB incrementa el riesgo de padecer DMT2 y enfermedades cardiovasculares. Según los criterios diagnósticos de la Organización mundial de la salud (OMS), se diagnostica en presencia de un IMC mayor a 30 kg/m². El sobrepeso se define según la OMS cuando el IMC está en el rango de 25 a 30 kg/m². La prevalencia global de la OB se ha duplicado desde el año 1980 al 2008 en el mundo, con un valor del 13% en el 2014 (Stevens et al 2012, <http://www.who.int/mediacentre/factsheets/fs311/es/>). A su vez, el sobrepeso ha pasado de ser de un 24.6 % en 1980, a un 39% en 2014. Los valores de prevalencia de la OB varían entre los distintos países y según el sexo (figuras 10 y 11). En España, la prevalencia de la OB es del 17.3% en población adulta (sin cambios en población masculina o femenina) según datos de la OMS del 2010, y para el sobrepeso, los porcentajes son del 57.9% en hombres y 49.8% en mujeres. En el estudio de nutrición y riesgo cardiovascular realizado en España (ENRICA), con un total de 12.883 individuos, se ha hallado una prevalencia de OB del 22.9% (24.4% en hombres y 21.4% para mujeres) (Gutiérrez-Fisac et al 2012). En otro estudio español, el ANIBES, se ha hallado una prevalencia de 19.9% para la obesidad, siendo un 22.7 % en hombres y un 17.3% en mujeres (López-Sobaler et al 2016). Además, se ha realizado una comparación de la prevalencia obesidad entre la zona sur de España (Andalucía), que presentaba un valor de 37.0%, y el resto del país, donde era un 26.6%, viéndose una diferencia significativa entre estas regiones (Valdés 2014).

La susceptibilidad de un individuo para presentar esta enfermedad está determinada por factores ambientales y genéticos. Un estilo de vida no saludable, caracterizado por un consumo excesivo de alimentos e inactividad física, es un factor de riesgo bien establecido para la aparición de la OB (Temelkova-Kurktschiev y Stefanov 2012).

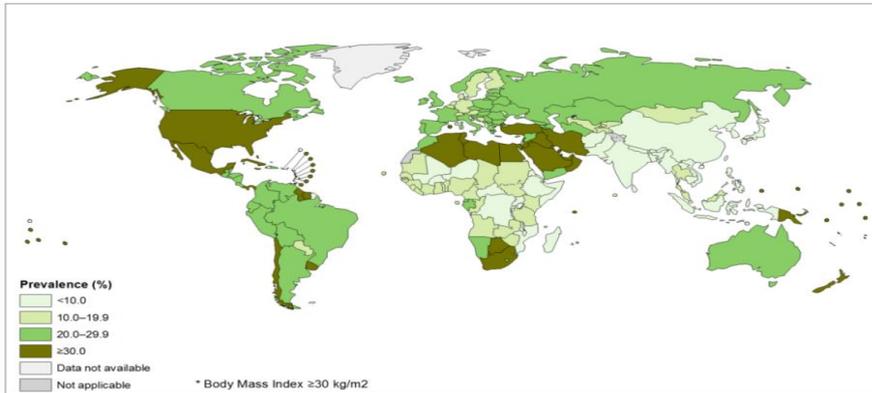


Figura 10. Prevalencia de la Obesidad por países, en hombres mayores de 18 años. (http://www.who.int/gho/ncd/risk_factors/overweight/en/)

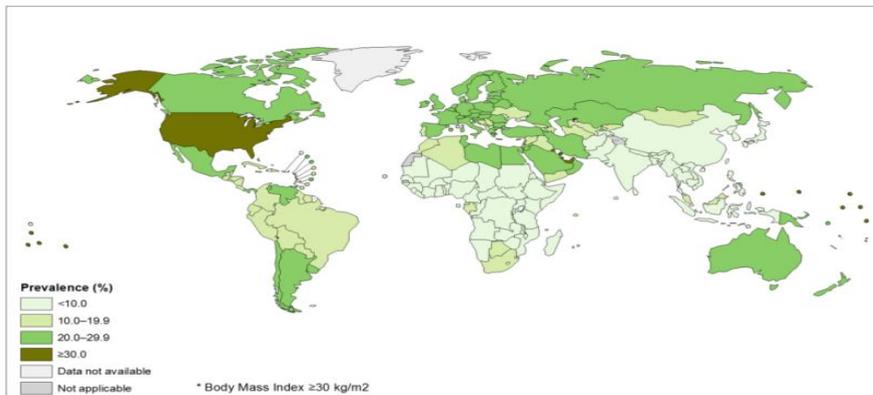


Figura 11. Prevalencia de la Obesidad por países, en mujeres mayores de 18 años. (http://www.who.int/gho/ncd/risk_factors/overweight/en/)

El fuerte impacto de los factores hereditarios sobre el desarrollo de la OB se ha confirmado a través de un gran número de estudios familiares, de adopción y en gemelos. En estudios de gemelos se ha demostrado que el 50-70 % de la variación del IMC puede ser explicado por la genética (Stunkard et al 1986, Allison et al 1996). Además, estudios de adopción y familiares han mostrado una heredabilidad de la OB en el rango del 20 a 60 % (Rice et al 1999, Stunkard et al 1986). Por otro lado, existen formas de obesidad mono-oligogenéticas. Los genes en los que se han encontrado mutaciones causales de OB son el de la leptina (LEP) (Montague et al 1997), Pro-opiomelanocortina (POMC) (Krude et al 1998), proproteína convertasa

subtilisina/kexina tipo 1 (PCSK1) (Jackson et al 1997), receptor de la melanocortina 4 (MC4R) (Farooqi et al 2003), factor neurotrófico derivado del cerebro (BDNF) (Gray et al 2006) y su receptor, tirosina quinasa B tipo 2 (NTRK2) (Yeo et al 2004). La mayoría de estos genes están involucrados en la regulación del comportamiento alimenticio (apetito), el balance energético y el peso corporal.

A partir de estudios de genes candidatos, se han hallado una gran cantidad de asociaciones entre genes y OB, pero sólo en algunos de ellos los resultados han sido replicados en otros estudios o se han hecho con tamaños muestrales adecuados (Karani et al 2010). Los genes cuyas variaciones se consideran en asociación fiable con OB son los siguientes:

- Adiponectina (ADIPOQ),
- Receptor β -adrenérgico 2 y 3 (ADRB2 y ADRB3),
- Proteína de unión al nucleótido guanina, polipéptido b3 (GNB3)
- Receptor c2 de la serotonina (HTR2C)
- Miembro 1 del receptor nuclear de la subfamilia 3, grupo C (NR3C1)
- Leptina (LEP)
- Receptor de la leptina (LEPR)
- receptor activado por el proliferador del peroxisoma gama (PPARG)
- Proteínas desacoplantes 1,2 y 3 (UCP1, UCP2 and UCP3)
- Receptor de la melanocortina 4 (MC4R)
- Factor neurotrófico derivado del cerebro (BDNF)
- Proproteína convertasa subtilisina/kexina tipo 1 (PCSK1)
- Receptor endocanabinoide 1 (CNR1)

Como en la mayoría de las enfermedades complejas, varios GWAS han ayudado a identificar una gran cantidad de variaciones asociadas con OB e IMC. Se han identificado múltiples loci en estos últimos años relacionados con ambos; como por ejemplo el del gen 2 inducido por insulina (INSIG2) y el gen FTO asociado con OB y masa grasa. Para el primero de ellos no se han obtenido a día de hoy resultados concluyentes. En cambio, FTO ha sido identificado como el gen más inequívocamente asociado con la susceptibilidad a la OB (Xia y Grant 2013).

En metaanálisis realizados por el Genetic Investigation of Anthropometric Traits (GIANT) consortium, se han descubierto muchos genes asociados con IMC y/o OB, tales como proteína transmembrana 18 (TMEM18), canal de potasio con dominio de

tetramerización 15 (KCTD15), glucosamina-6-fosfato desaminasa 2 (GNPDA2), proteína 1B con dominio SH2 (SH2B1), transportador mitocondrial 2 (MTCH2) y el regulador de crecimiento neuronal 1 (NEGR1). En la web del EMBL-EBI se muestra un catálogo de los estudios GWAS recopilados hasta el momento para la obesidad, y todos sus hallazgos (<http://www.ebi.ac.uk/gwas/search?query=obesity>). Se han hecho cálculos basados en los resultados de los GWAS que indican que todavía faltan por identificar alrededor de 250 variantes frecuentes, con un efecto similar a los SNPs hallados en los GWAS (Speliotes et al 2010). El reto actual está además en entender los mecanismos biológicos de los loci asociados y cómo confieren el riesgo para la OB.

Otro aspecto importante a tener en cuenta es cómo influye la epigenética en el riesgo de padecer enfermedades complejas como la OB y la diabetes. Se han empezado a estudiar tres marcadores epigenéticos principales, y los mecanismos que pueden estar actuando en el tejido adiposo y en la OB (Marti y Ordavas 2011). Estos marcadores son la metilación del genoma, cambios en la organización cromatínica a través de modificación de histonas, y los ARN no codificantes. Así, diferentes genes implicados en procesos relacionados con la obesidad son regulados por la metilación del promotor. Como ejemplo, el gen PPAR2 es desmetilado progresivamente en la diferenciación del adipocito, lo que provee evidencia de que la metilación del ADN influye en la expresión de genes específicos del adipocito (Fujiki et al 2009). Otras técnicas, como el estudio de la secuencia del genoma completo tratado con bisulfito, o la inmunoprecipitación cromatínica seguida de secuenciación masiva (ChIP-Seq), serán muy útiles para encontrar más detalles acerca de la metilación en relación con enfermedades como la OB.

Existe una relación importante entre una gran cantidad de loci de susceptibilidad a la OB y el riesgo de padecer DMT2. El locus FTO se identificó por primera vez en un GWAS de DMT2, pero al ajustar los resultados según el IMC la asociación hallada se perdió, lo que indicaba que el IMC mediaba la asociación de ese gen y la DMT2 (Frayling et al 2007). Otros loci que confieren susceptibilidad para la OB, como son TMEM18, GNPDA2, NEGR1, ETV5, DGKG, BCDIN3D y FAIM2, también presentan este comportamiento, muestran una asociación evidente con la DMT2

pero el IMC parece jugar un papel importante en la aparición de esa asociación (Thorleifsson et al y Willer et al 2009).

1.2.3.2.2 Obesidad central

Además de la OB general indicada anteriormente, la distribución de la grasa corporal también influye en el riesgo de DMT2 y de otros fenotipos relacionados con ella. El parámetro más empleado para caracterizarla es la circunferencia de la cintura. Según la IDF, se define la obesidad central o abdominal (OB Central) cuando la cintura es mayor o igual a 94 cm en hombres y a 80 cm en mujeres. La Adult Treatment Panel III (ATP III) sin embargo, recomienda unos valores de corte de cintura de 102 cm en hombres y de 88 cm en mujeres. La presencia de OB Central es uno de los componentes principales para diagnosticar el síndrome metabólico (http://www.idf.org/webdata/docs/MetSyndrome_FINAL.pdf).

La OMS ha publicado algunos estudios de la prevalencia de la OB Central, como por ejemplo el realizado en Grecia, donde se hallaron un 54.3% de hombres y un 56.5% de mujeres que presentaban OB Central, utilizando los valores de corte de la IDF (http://www.euro.who.int/_data/assets/pdf_file/0010/74746/E90711.pdf). En Inglaterra, considerando los criterios de la ATP III, se vio una prevalencia del 41,1% en mujeres y del 31,1% en hombres. En España, en el estudio ANIBES de 2016, se encontró que el 24.6% de los hombres, y un 31.4 % de las mujeres tienen OB Central (López-Sobaler et al 2016). En la figura 12 se muestran las prevalencias de OB Central ajustada por edad en España (Gutiérrez-Fisac et al 2012).

1.2.3.2.3 Hipertensión arterial

Cerca de mil millones de personas en todo el mundo padecen hipertensión arterial (HTA), con una prevalencia global del 22%. Esta enfermedad se diagnostica cuando la presión arterial sistólica (PAS) es mayor a 140 mm Hg y/o la presión arterial diastólica (PAD) es mayor a 90 mm Hg (International Consortium for Blood Pressure Genome-Wide Association Studies et al 2011). La presión arterial se ve afectada por factores ambientales y adquiridos, como son la edad, el sexo, el IMC, la tolerancia a la glucosa alterada y el consumo de tabaco, y por factores genéticos.

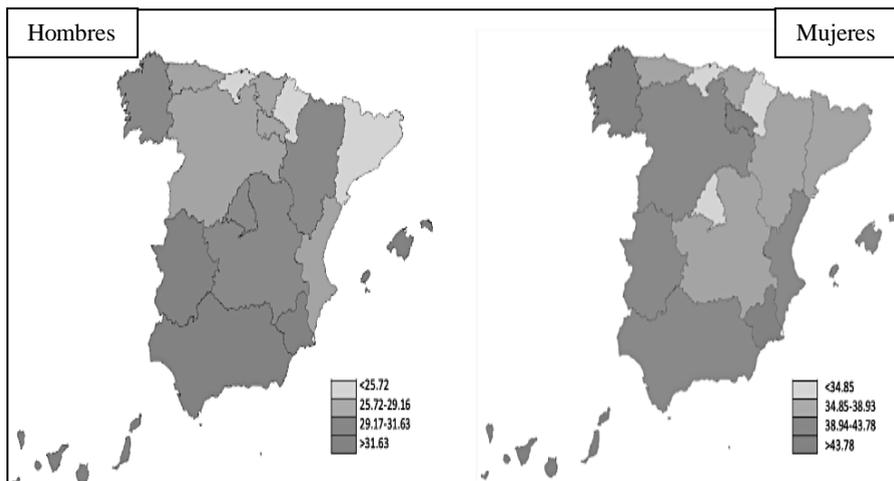


Figura 12. Variación regional de los porcentajes de prevalencia de la obesidad abdominal ajustada por edad en España, para hombres y mujeres, período 2008-2010 (Obes Rev. 2012; 13(4):388-92).

Las prevalencias en hombres y mujeres obtenidas por la OMS varían según los países (Figuras 13 y 14). La prevalencia de hipertensión en mujeres es menor que en hombres hasta los 55-64 años, edad a partir de la cual se equipara con la de los hombres, y empieza a aumentar a mayor velocidad. En la población adulta española, la prevalencia hallada a través del estudio Di@bet.es fue del 42.6% (Menéndez et al 2016), siendo mayor en hombres (49.9%) que en mujeres (37.1%).

La HTA está presente en más de un 60 % de los pacientes con DMT2, siendo un factor de riesgo para esta enfermedad. La prevalencia de la hipertensión en pacientes diabéticos ha sido descrita en varios estudios realizados en distintas regiones del mundo (Colosia et al 2013). En pacientes diabéticos, las mujeres alcanzan la prevalencia del hombre a la edad de 45-54 años, una década antes con respecto a los no diabéticos. Su desarrollo puede preceder y predecir la aparición de la hiperglucemia. Estas dos patologías afectan en general a los mismos órganos, principalmente al árbol vascular (Grossman y Messerli 2008). La presencia simultánea de DMT2 e hipertensión acelera el desarrollo de la retinopatía diabética y la aparición de enfermedades cerebrales, así como el riesgo cardiovascular en general y la disminución de la función renal.

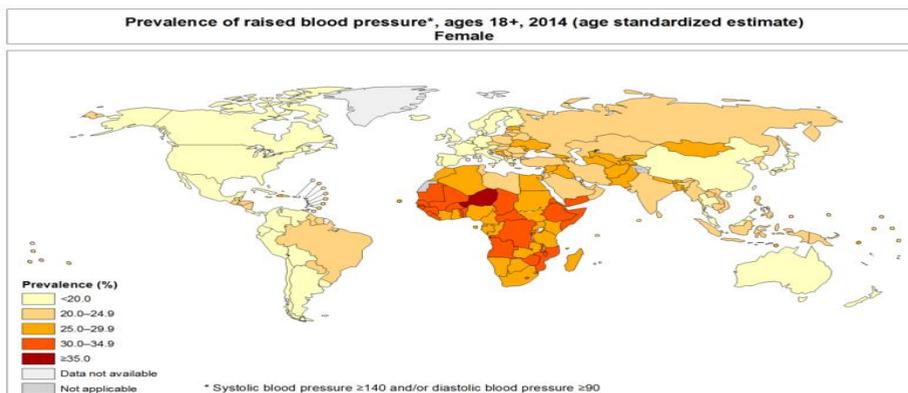


Figura 13: Prevalencias de presión arterial alta en mujeres mayores de 18 años. Datos obtenidos de la OMS 2014 (<http://gamapserver.who.int/mapLibrary/app/searchResults.aspx>).

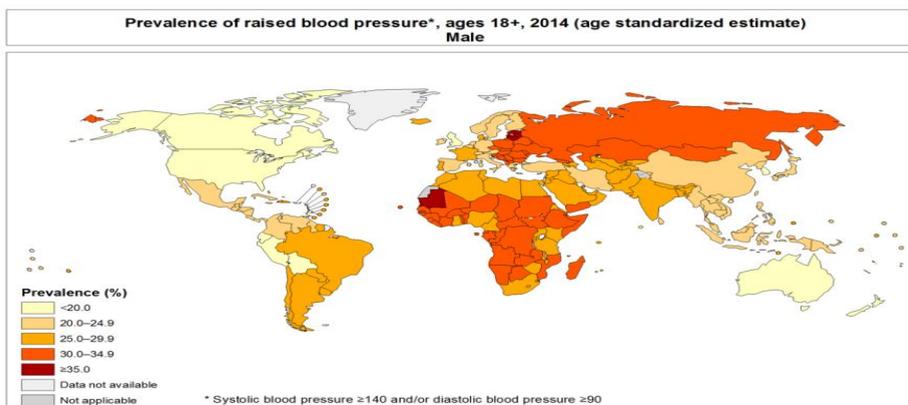


Figura 14: Prevalencias de presión arterial alta en hombres mayores de 18 años. Datos obtenidos de la OMS 2014 (<http://gamapserver.who.int/mapLibrary/app/searchResults.aspx>).

La tolerancia reducida a la glucosa y la presión arterial elevada están fuertemente relacionadas, siendo esta última más frecuente en pacientes diabéticos (Ferrannini y Cushman 2012). Un defecto compartido entre la DMT2 y la hipertensión esencial o primaria es la insulino-resistencia (Ferrannini et al 1987), que no está descrita en pacientes con hipertensión secundaria a otras patologías. La posible relación entre la insulina y la presión arterial se podría explicar por las acciones vasculares de esta hormona, que se resumen en la tabla 2. Por su parte, la OB es un potente factor ambiental que influye tanto en el metabolismo glucídico como en la presión arterial.

| | Acciones potencialmente positivas de la Insulina | Acciones potencialmente negativas de la Insulina |
|---------------------------|----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Mecanismo | Estimulación de la óxido nítrico sintasa; inhibición del factor nuclear kapa beta; inhibición de la agregación plaquetaria | Hipertrofia vía la proteína quinasa activada por mitógeno; estimulación de la endotelina1; hiperpolarización de membrana; activación adrenérgica; antinatriuresis; estimulación del inhibidor tipo1 del activador del plasminógeno; especies reactivas del oxígeno |
| Efecto fisiológico | Vasodilatación; antiinflamatoria; antitrombótica | Proliferación de las células musculares lisas; vasoconstricción; intervalo QTc extendido; retención de fluidos; protrombótica |

Tabla 2. Acciones vasculares de la insulina (Lancet 2012; 380(9841): 601-10)

Diferentes estudios han mostrado parte de la compleja arquitectura genética de esta patología poligénica (Simino et al 2012). Se ha asociado la variación de la presión arterial con genes tales como los que codifican para canales iónicos o para reguladores de estos canales, para la ruta de señalización del sistema aldosterona, para la vasoconstricción, y para la inflamación, en estudios de genes candidatos. También se han encontrado asociaciones entre varios genes codificantes para enzimas antioxidantes (SOD3, CAT, GPX1, TXN) y la presión arterial, y el riesgo de padecer hipertensión (Mansego et al 2011). Más de cincuenta loci a lo largo de todo el genoma se han asociado con la presión arterial y la hipertensión en estudios GWAS (Simino et al 2012), Estos se detallan en la figura 15. Actualmente se están desarrollando estudios mediante análisis del exoma para identificar variaciones involucradas en fenotipos específicos 1 (base de datos RePORTER del National Institute of Health, NIH). Un ejemplo podría ser el hallazgo de mutaciones en los genes kelch-like 3 (KLHL3) y cullina 3 (CUL3) en un estudio del desorden Mendeliano llamado Pseudohipoaldosteronismo tipo 2 (pacientes que presentan hipertensión esencial, hipercalemia y acidosis metabólica).

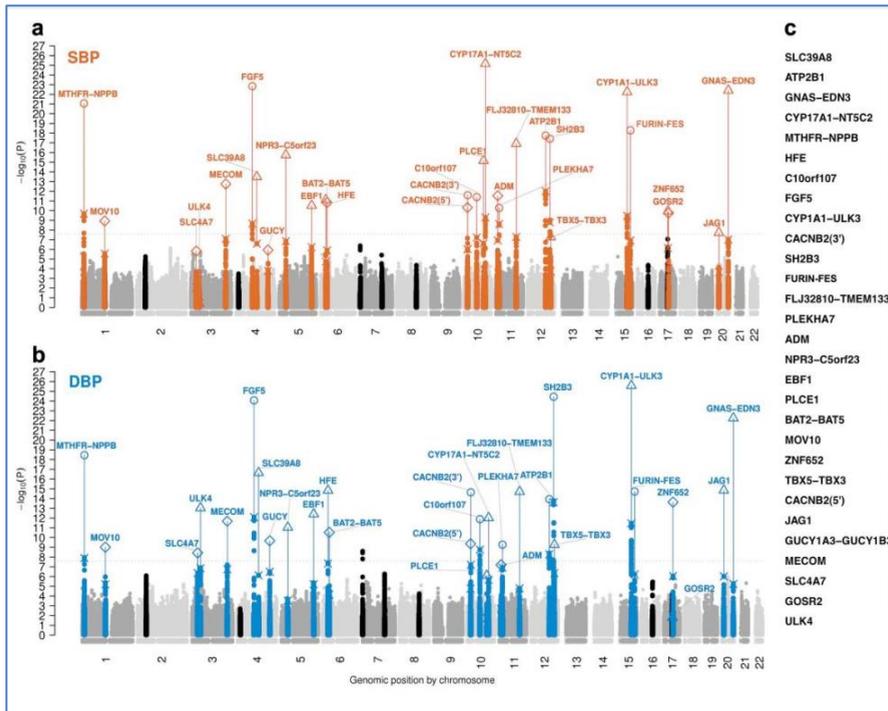


Figura 15. Manhattan plot de GWAS para PAS (a) y PAD (b). Los SNPs asociados significativamente se muestran en naranja para PAS y en azul para PAD. El listado de los polimorfismos asociados con ambos parámetros se muestra a la derecha (c) (Nature. 2011; 478(7367):103-9).

A nivel epigenético, se han descrito miARN que actúan sobre el sistema aldosterona-angiotensina-renina, sobre las células vasculares del músculo liso o sobre genes de otras rutas involucradas en la hipertensión esencial (Batkai y Thum 2012). Con respecto al estudio de la metilación, en animales se han encontrado modificaciones de la metilación (Rivière et al 2011), y en las histonas del gen de la enzima convertidora de la angiotensina 1(ACE) relacionadas con esta enfermedad (Lee et al 2012). Su relación con la disfunción mitocondrial se ha demostrado en varios trabajos, que vinculan al sistema renina-angiotensina y la deficiencia de vitamina D (Manucha et al 2015) y a los productos ROS (Dikalov y Dikalova 2016) con la hipertensión.

1.2.4 Tratamiento de la DMT2

La diabetes requiere un tratamiento médico continuo que va más allá del control glucémico, debiéndose reducir al máximo los factores de riesgo ambientales para minimizar las complicaciones a largo plazo de la enfermedad. La educación de cómo debe manejarse el paciente en el día a día, y un fuerte apoyo psicológico son críticos para prevenir episodios agudos y también las complicaciones futuras. Un buen cribado, diagnóstico y acciones terapéuticas en el momento adecuado, se sabe que afectan favorablemente la evolución del paciente diabético, por lo cual estos pasos son muy importantes. Instaurar el tratamiento cuando antes es esencial para obtener unos mejores resultados.

Hay ciertas entidades que realizan guías para orientar a pacientes, médicos e investigadores acerca de la diabetes. Elaboran un resumen del manejo de la diabetes con las prácticas de mayor consenso en el mundo y dan herramientas para evaluar la calidad del tratamiento recibido. La American Diabetes Association (ADA) ha publicado la última edición en 2015 de estas recomendaciones, que se pueden modificar en función del criterio médico para cada caso en particular (ADA 2015).

Se recomienda realizar una evaluación médica completa del paciente en su visita inicial, para clasificar su diabetes y detectar complicaciones. Si el paciente ya tiene establecida la diabetes, es preciso revisar previos tratamientos, comprobar si están controlados los factores de riesgo y ver el nivel de daño orgánico.

Lo primero que se debe modificar en estos pacientes son hábitos como la nutrición, la actividad física, el tabaquismo, dar apoyo psicosocial y evaluar la inmunización. Para muchos diabéticos, la parte más difícil del tratamiento es la dieta, que debe ajustarse según las necesidades de cada paciente, teniendo en cuenta los siguientes objetivos:

- Alcanzar los niveles recomendados de glucemia, de lípidos en sangre, y de presión arterial
- Mantener un peso corporal adecuado
- Retrasar y prevenir las complicaciones de la diabetes

Los tratamientos de las 2 formas más comunes de la diabetes, la tipo 1 y 2, son muy diferentes, debido a que las causas que producen la enfermedad difieren. En la diabetes mellitus tipo 1, al producirse una destrucción de las células beta del páncreas, la terapia de elección es la de sustitución de la insulina, que está en déficit. Estos pacientes deben ser enseñados a compatibilizar la dosis de insulina prandial con la ingesta de carbohidratos, la glucemia preprandial y la actividad física. Es preferible el uso de análogos de la insulina, especialmente en casos de alto riesgo de hipoglucemia. En general deben administrarse múltiples dosis de insulina, tres o cuatro veces al día.

En el caso de la DMT2, el punto de mira del tratamiento es la reducción de la hiperglucemia y controlar las enfermedades concomitantes, especialmente las cardiovasculares (hipertensión y dislipemias) y el daño orgánico. Un cambio en el estilo de vida, dieta y ejercicio, y una terapia farmacológica intensiva para bajar la glucemia y para reducir factores de riesgo, disminuye los eventos cardiovasculares, las tasas de muerte y la progresión a una enfermedad renal de estado terminal, como así también la retinopatía diabética. Se utilizan agentes farmacológicos como la metformina, inhibidores de la alfa glucosidasa, orlistat y tiazolidinedionas. Con el primero se ha demostrado muy fehacientemente que es un fármaco que previene la diabetes. En la figura 16 se muestra el algoritmo que recomienda seguir en los pacientes con DMT2 la IDF (<https://www.idf.org/e-library/guidelines.html>).

La hiperglucemia debe ser controlada a largo plazo a través de la medida de la hemoglobina glicosilada (HbA1C) por lo menos 2 veces al año, en pacientes con diabetes. El manejo de la hiperglucemia debe ser individualizado de acuerdo a la duración en el paciente de la DMT2, de la esperanza de vida, de la presencia de enfermedades asociadas o complicaciones vasculares, del potencial para hacer hipoglucemias, y otras características específicas de cada paciente que puedan afectar a la terapia.

La ADA y la EASD recomiendan unos valores de HbA1C de 7 % en los pacientes diabéticos, y la American association of clinical Endocrinologists (AACE) de 6,5%.

IDF Treatment Algorithm for People with Type 2 Diabetes

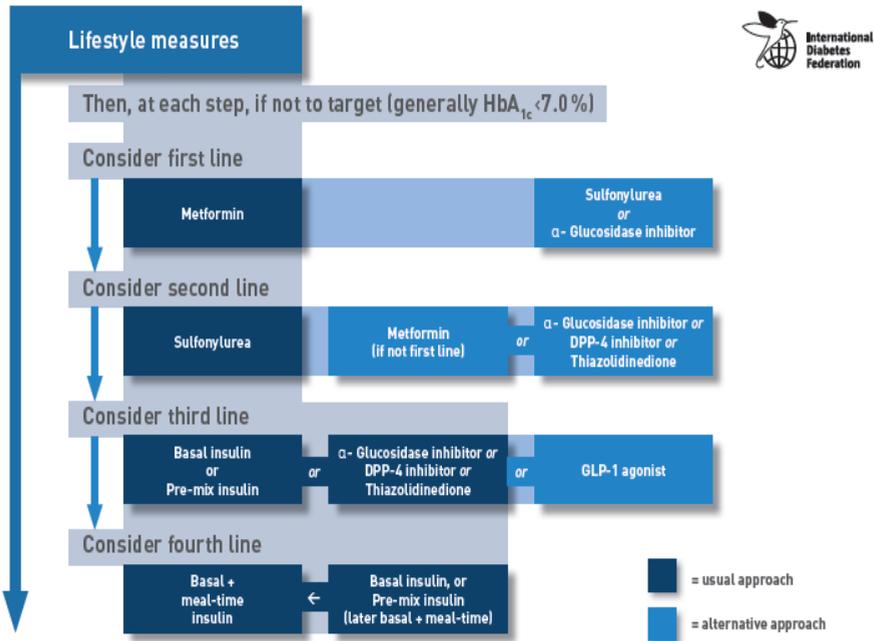


Figura 16. Recomendaciones para el tratamiento de la DMT2 según la IDF, obtenidas del Clinical Guidelines Task Force, Global Guideline for Type 2 Diabetes del 2012 (<https://www.idf.org/e-library/guidelines.html>).

Existen nuevas terapias que están siendo evaluadas como posibles opciones de tratamiento (Cornell 2015), como el tratamiento con los agonistas de receptores, como los del receptor acoplado a proteína G119 (GPR119). Varios agonistas de GPR119 ya están en desarrollo como fármacos para la DMT2 (GlaxoSmithKline 2015). Este receptor se expresa predominantemente en las células beta del páncreas y tracto gastrointestinal en humanos y su estimulación lleva a un aumento de la liberación de insulina, GLP-1 y GIP. El receptor 1 de ácidos grasos libres (FFAR1) se expresa principalmente en las células beta del páncreas y su activación provoca una secreción de insulina inducida por ácidos grasos libres, y estimulada por glucosa. El activador de FFAR1 TAK-875 ha reducido significativamente la HbA1C en un ensayo clínico (Burant et al 2012).

Con respecto a la medicina personalizada en la diabetes mellitus, todavía se sigue necesitando determinar la terapia apropiada para cada individuo, ya que un número significativo de tratamientos en monoterapia fallan después de 3 años, y la morbilidad y mortalidad continúan su curso en el paciente (Kleinberger y Pollin 2015). Por otro lado, las terapias inapropiadas pueden llevar al paciente a aumentar el riesgo de eventos hipoglucémicos. El tratamiento específico para cada enfermo puede orientarse según su genética, lo que podría, en muchos casos, disminuir el número de fallos terapéuticos y reducir las complicaciones asociadas a la diabetes. Algunas formas monogénicas de diabetes como la MODY, presentan desde hace un tiempo la oportunidad de tener una medicina personalizada. Saber si es MODY 1 o 3, es decir, hacer el diagnóstico genético, ayuda a ajustar la dosis de la sulfonilurea a utilizar, debido a que los pacientes que presentan una de estas 2 formas son hipersensibles a este medicamento. En la diabetes mellitus neonatal (DMN), hay una disminución en la secreción de insulina debido más comúnmente a una mutación en el gen *KCNJ11* o *ABCC8*. Muchos de estos bebés pueden ser tratados satisfactoriamente con altas dosis de sulfonilureas y no con el tratamiento por defecto para la diabetes neonatal, que es la insulina. Así, los pacientes reciben una terapia menos invasiva, más barata y más efectiva, poniéndose en menor riesgo de hipoglucemia.

Existen ya numerosos estudios que permitirían mejorar el tratamiento en base a la farmacogenética en la DMT2 (<https://www.pharmgkb.org/>). Sin embargo, la metformina es el fármaco más ampliamente utilizado en estos pacientes a pesar de su alta variabilidad de eficacia entre ellos. Aunque su mecanismo de acción no está del todo esclarecido, se ha identificado un SNP cercano al gen de la ATM serinotreonina quinasa (ATM), que se asocia con una reducción de la HbA1c en respuesta a la metformina (van Leeuwen et al 2012). Se ha visto que múltiples variantes genéticas de un gen transportador, el miembro 22 de la familia de transportadores de soluto (*SLC22A1*), se asocian con una eficacia disminuida de este medicamento (Tzvetkov et al 2009). Estos resultados podrían ser usados para predecir la eficacia de esta droga en los pacientes antes de iniciar el tratamiento. Otros estudios han encontrado asociaciones entre las sulfonilureas y la respuesta al tratamiento; en este

sentido polimorfismos en los genes *KCNJ11* y *ABCC8* se han asociado con alteraciones en la eficacia de las sulfonilureas, y un haplotipo de estos genes ha demostrado ser menos sensible a la inhibición de la sulfonilurea en análisis de patch-clamp (Becker et al 2013). Genes tan importantes en la DMT2 como *TCF7L2* e *IRS-1* también se han relacionado con diferencias en la respuesta a la administración de esta medicación. Además, se han hallado variaciones en las enzimas que metabolizan la sulfonilurea que afectan a la eficacia de la droga, como variaciones en el gen *CYP2C9* (Holstein et al 2005). Del grupo de las tiazolidinedionas, agonistas del *PPARG*, también se han encontrado variantes genéticas con mayor predisposición a efectos secundarios adversos. Otra medicación antidiabética oral de la que se han relacionado variaciones génicas con su eficacia, es la que actúa a través de la ruta de señalización de las incretinas, que incluye a los análogos de las hormonas intestinales péptido 1 glucagón like (*GLP-1*) y péptido insulino-trópico dependiente de glucosa (*GIP*), y a los inhibidores de la dipeptidil-peptidasa 4 (*DPP-4*) (Jamaluddin et al 2014). Recientemente, los inhibidores del cotransportador sodio-glucosa tipo 2 del túbulo proximal renal (*SGLT2*) han mostrado su eficacia en el tratamiento de la DMT2 (Chao y Henry 2010, Fioretto et al 2015). El mecanismo de acción de estos inhibidores es impedir la reabsorción de glucosa en el riñón, aumentando así su excreción urinaria. El atractivo de estos fármacos es que actúan de forma independiente de la insulina, pero su inconveniente es que dependen de la filtración de glucosa en el riñón. Cuando la velocidad de filtración glomerular decae, se reduce la cantidad de glucosa filtrada, disminuyendo la glucosuria y su eficacia. La farmacogenómica en diabetes ha sido poco desarrollada hasta el momento.

1.3 Cadena mitocondrial respiratoria

1.3.1 Características generales y mecanismo molecular

La cadena mitocondrial respiratoria (CMR) es un sistema biológico esencial en el metabolismo energético. Está involucrada en la producción del ATP celular mediante la fosforilación oxidativa (OXPHOS), que realiza a partir de productos del metabolismo de hidratos de carbono y lípidos. La CMR produce un gradiente de

protones necesario para la síntesis de ATP. Este gradiente se genera a través de cuatro complejos enzimáticos: I, II, III y IV. Estos transportan electrones de moléculas donantes, como el NADH en el complejo I y FADH₂ para el II, hacia los aceptores de electrones. Finalmente, en el complejo IV los electrones llegan a su último aceptor que es el oxígeno molecular (Johannsen et al 2009). Todo este proceso se realiza en la membrana interna mitocondrial.

La glucosa es metabolizada a piruvato y después a acetil-CoA y CO₂ en la glicólisis, generando NADH + H⁺. Por otro lado, los ácidos grasos son oxidados, dentro de la mitocondria, a través de la β-oxidación, para producir también acetil-CoA y NADH + H⁺, además de FADH₂. El acetil-CoA entra finalmente en el ciclo del ácido tricarbóxico (TCA), en donde se obtienen los hidrógenos de los hidrocarbónos para generar NADH + H⁺. Dos electrones son transportados desde cada molécula de NADH + H⁺ hacia la NADH Deshidrogenasa (complejo I de la CMR) o desde enzimas que transportan FADH₂ como la succinato deshidrogenasa (complejo II de la CMR), para reducir la ubiquinona (coenzima Q10, CoQ) a ubiquinol (CoQH₂). Los electrones van de manera sucesiva desde el CoQH₂ al complejo III (complejo bc1), citocromo c, citocromo c oxidasa (complejo IV, COX), y finalmente al oxígeno para dar H₂O (figura 17, A). La energía liberada por este flujo de electrones es utilizada por los complejos I, III y IV para bombear protones a través de la membrana interna mitocondrial, creando así un gradiente electroquímico de protones (Wallace y Fan 2010). La energía almacenada en este gradiente es utilizada para múltiples propósitos como para importar proteínas y calcio hacia la mitocondria, para generar calor, y sobretodo para sintetizar ATP dentro de la matriz mitocondrial. La ATP sintetasa (complejo V) usa la energía producida por el flujo de protones hacia la matriz para convertir el ADP y Pi en ATP. Finalmente, el ATP de la matriz se intercambia por ADP citosólico a través de traslocadores de nucleótidos de adenina presentes en la membrana interna mitocondrial (Wallace 2007).

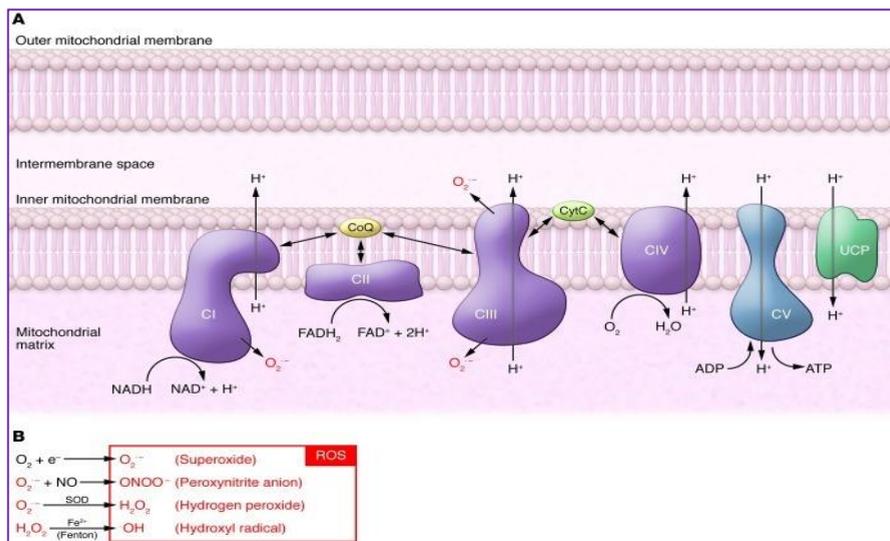


Figura 17. Esquema del sistema de fosforilación oxidativa (A) y la producción de ROS (B) de la CMR. (J Clin Invest. 2013; 123(3):951-7).

Además de su papel en la producción de energía, la mitocondria es el principal productor de especies reactivas del oxígeno (ROS) en la célula. Estas moléculas están representadas fundamentalmente por el anión superóxido (figura 17, B). Dentro de la cadena respiratoria, los complejos I y III son sus mayores productores. (Kadenbach 2012). Se han implicado a los ROS mitocondriales en numerosas enfermedades complejas, como las neurodegenerativas (Lin and Beal 2006), y en el proceso de envejecimiento (Muller et al. 2007; Navarro y Boveris 2007; Bratic y Larsson 2013). Hallazgos recientes sugieren que los ROS mitocondriales pueden ser a su vez, componentes integrales de la señalización celular (Vandecasteele et al 2001; Logan y Knight, 2003) y de la muerte celular programada (Jones 2000; Youle y Karbowski, 2005).

La mitocondria también actúa en otros procesos metabólicos, incluidos la biosíntesis de aminoácidos, cofactores de vitaminas y ácidos grasos (Mackenzie y McIntosh, 1999; Bowsher y Tobin, 2001), y es uno de los tres compartimentos involucrados en la fotorespiración (Douce and Neuburger, 1999).

La mitocondria posee de 2 a 10 copias de ADN circular de doble cadena que es heredado de forma materna (Majamaa et al 1998). A pesar de contener su propio

genoma, su sistema de transcripción y su maquinaria de síntesis de proteínas, la mayoría de los complejos enzimáticos de la CMR poseen polipéptidos que están codificados por el genoma nuclear, sintetizados en el citosol e importados a la mitocondria (Logan 2006). El ADN mitocondrial (mtADN) de los mamíferos codifica para 2 ARNs ribosomales, 22 ARNs de transferencia (tRNA) y 13 proteínas. Estas últimas constituyen algunas subunidades de los complejos de la CMR (7 del complejo I, 1 del complejo III, 3 del complejo IV, y 2 del complejo V). La transcripción y la traducción mitocondrial difieren del mecanismo nuclear y requieren más de 100 proteínas codificadas por el núcleo para poder llevarse a cabo. A estas 100 proteínas hay que sumar otras 1000 sintetizadas a partir de genes nucleares. Todas ellas son necesarias para sintetizar toda la maquinaria mitocondrial, incluyendo los complejos de la cadena respiratoria. En la figura 18 se muestran algunas de las proteínas involucradas en procesos mitocondriales. Las proteínas nucleares deben ser transportadas hacia la mitocondria por un sistema de transporte de proteínas muy sofisticado (Schmidt et al. 2010).

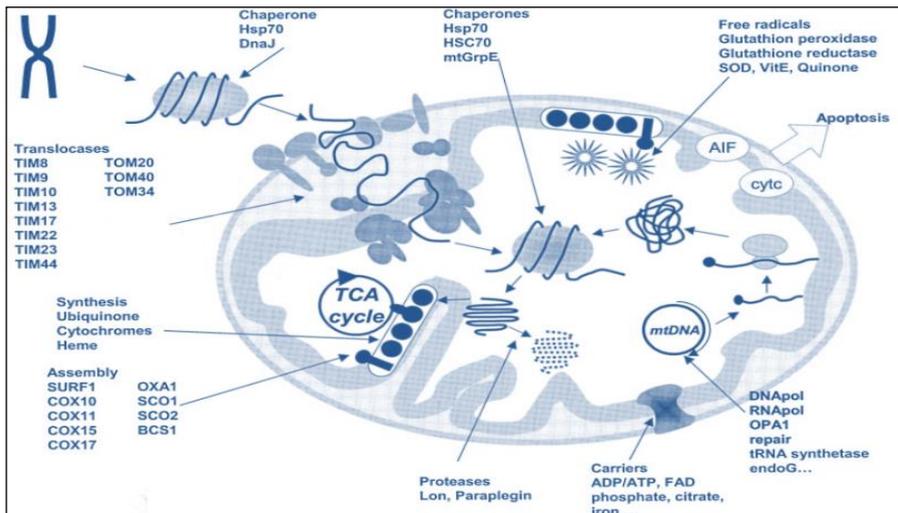


Figura 18. Esquema de algunas de las proteínas que actúan en la mitocondria. SOD: superóxido dismutasa; cytc: citocromo c; ADNpol: ADN polimerasa; RNAPol: ARN polimerasa; VitE: vitamina E; AIF: factor de inducción de apoptosis (J Am Soc Nephrol. 2003; 14(12):2995-3007).

1.3.2 Cadena mitocondrial respiratoria y su relación con la DMT2

En estas últimas décadas, la deficiencia en el metabolismo energético ha surgido como una alternativa a la explicación de la etiología de algunas enfermedades complejas. La OXPHOS provee energía a la mayoría de los tejidos y órganos, por lo que una deficiencia de la CMR puede generar un amplio espectro de sintomatología y desórdenes. Alteraciones del ADN mitocondrial y de genes implicados en el funcionamiento de la mitocondria, involucran a este orgánulo en un amplio espectro de enfermedades metabólicas y degenerativas (Fosslien 2001; Munnich y Rustin 2001; Rötig y Munnich 2003; Wallace y Fan 2010). Mutaciones en genes nucleares que codifican para subunidades de la CMR son causales de enfermedades tales como el síndrome de Leigh o el paraganglioma (Zhu et al 2009). Algunas de estas mutaciones inactivan los genes que intervienen en la OXPHOS (Procaccio y Wallace 2004; Zhu et al 1998), desestabilizan al mtADN (Carrozzo et al 2007; Kaukonen et al 2001; Palmieri et al 2005; Saada et al 2001) o perturban la fusión y fisión mitocondrial (Delettre et al 2000 y Zuchner et al 2004; Zhu et al 2009). Ejemplos de estas variaciones y las enfermedades que producen se indican en la tabla 3 (Rotig y Munnich 2003). En los años 90, se incluyó a la diabetes como uno de los fenotipos de la enfermedad mitocondrial (Barrett 2001). En ciertos casos, la diabetes es producida por mutaciones en el mtADN, pero también puede estar asociada a desórdenes mitocondriales, como la Ataxia de Friederich o la enfermedad de Huntington (Schapira 2012), o como parte de una encefalopatía (MELAS).

La mitocondria juega un papel crucial en el mecanismo de señalización de la glucosa implicado en la liberación de insulina. Muchos defectos mitocondriales afectan a este proceso, provocando una hiperglucemia persistente y una reducción en la liberación de la insulina (Jitrapakdee et al 2010 y Barrett 2001). Los mecanismos que se han asociado a la patogenia de la DMT2 incluyen una inadecuada función o capacidad mitocondrial, una alteración en la señalización de liberación de insulina debida a acumulación lipídica celular, a señales proinflamatorias y a estrés del retículo endoplásmico, y una reducida secreción celular de insulina (Patti y Corvera 2010). La secreción de esta hormona en respuesta a la elevación de la glucemia post-prandial depende de una secuencia de eventos metabólicos, que incluyen el

transporte de la glucosa a través del transportador GLUT2, la fosforilación de la glucosa por la glucoquinasa, la producción de NADH y piruvato en la glicólisis, y la estimulación de la OXPHOS mitocondrial. En este último paso, el ATP producido se eleva, lo que lleva al cierre de los canales de potasio ATP-dependientes. Entonces, la membrana se despolariza y se abren los canales de calcio, provocando un aumento intracelular de los niveles de este elemento que culmina en la secreción de la insulina. Si la disponibilidad de ATP es baja en las células beta, esto puede comprometer a toda la ruta de secreción de insulina en respuesta a las glucemias elevadas (Gerbitz et al 1996).

| Gene | Function | Enzyme Deficiency | Clinical Features |
|------------------------------------------|--------------------------------|--------------------------------------|------------------------------------------------------|
| Structural RC genes | | | |
| <i>FpSDH</i> | Complex II subunit | SDH (complex II) | Leigh syndrome |
| <i>NDUFV1</i> | Complex I subunit | Complex I | Leukodystrophy and myoclonic epilepsy |
| <i>NDUFSS</i> | Complex I subunit | Complex I | Leigh syndrome |
| <i>AQDQ</i> | Complex I subunit | Complex I | Multivisceral involvement |
| <i>NDUFS7</i> | Complex I subunit | Complex I | Leigh syndrome |
| <i>NDUFS1</i> | Complex I subunit | Complex I | Leigh syndrome |
| <i>NDUFV2</i> | Complex I subunit | Complex I | Cardiomyopathy, encephalopathy |
| Genes involved in RC assembly | | | |
| <i>SURF1</i> | COX assembly | COX | Leigh syndrome |
| <i>COX10</i> | Heme maturation | COX | Tubulopathy and leukodystrophy |
| <i>SCO1</i> | Copper metabolism | COX | Hepatopathy and ketoacidotic coma |
| <i>SCO2</i> | Copper metabolism | COX | Cardiomyopathy |
| <i>COX15</i> | Heme maturation | COX | Cardiomyopathy |
| <i>LRPPRC</i> | | COX | Leigh syndrome |
| <i>BCS1</i> | Complex III assembly | Complex III | Tubulopathy and liver failure, GRACILE syndrome |
| <i>HSP60</i> | Chaperone | | Spastic paraplegia (SPG13) |
| <i>OPA1</i> | Dynamin-related protein | | Dominant optic atrophy |
| <i>ABC7</i> | Iron carrier | | Sideroblastic anemia |
| frataxine | Fe-S cluster assembly | RC deficiency | Friedreich's ataxia |
| <i>WD</i> | Copper-transporting ATPase | | Wilson's disease |
| paraplegin | Chaperone-like protein | Multiple deficiency | Spastic paraplegia |
| <i>DDP</i> | Mitochondrial import machinery | | Deafness, dystonia, mental deficiency, and blindness |
| Genes involved in mtDNA stability | | | |
| <i>TP</i> | Thymidine phosphorylase | Multiple mtDNA deletion | MNGIE |
| <i>G4.5</i> | Tafazzin | Multiple RC deficiency | Barth syndrome |
| Twinkle | | Multiple deficiency, mtDNA deletions | PEO |
| <i>ANT1</i> | ADP/ATP translocator | | PEO |
| <i>POLG</i> | Polymerase γ | | PEO |
| <i>TK2</i> | Thymidine kinase | mtDNA depletion | |
| <i>DGUOK</i> | Deoxyguanosine kinase | mtDNA depletion | |

Tabla 3. Genes nucleares implicados en desórdenes mitocondriales. SDH: Succinato deshidrogenasa; GRACILE: retardo de crecimiento, aminoaciduria, colestasis, sobrecarga de hierro, lactoacidosis y muerte temprana; PEO: oftalmoplegia externa progresiva; MNGIE: encefalopatía neurogastrointestinal mitocondrial; SPG13: paraplejía espástica; COX: citocromo c oxidasa (J Am Soc Nephrol 2003; 14: 2995–3007).

En este sentido, se ha encontrado un metabolismo anormal de la mitocondria, con una masa mitocondrial reducida, presente en biopsias musculares e in vivo, en pacientes obesos y con DMT2 (Ritov et al 2010). La reducción de la masa mitocondrial se ha asociado con una disminución en un 20-30 % en la expresión de los genes nucleares codificantes de subunidades de la OXPHOS (Patti et al 2003). Hallazgos similares se han encontrado en las mitocondrias de los tejidos adiposo y hepático en pacientes (Martin y McGee 2013).

La disfunción mitocondrial y la alteración de los genes involucrados en la OXPHOS están relacionadas también con la RI (Parish y Petersen 2005; Abdul-Ghani y DeFronzo 2008). Existen múltiples causas que pueden provocar un mal funcionamiento mitocondrial (figura 19). Mutaciones en los genes mitocondriales causadas por el envejecimiento o condiciones de estrés celular (sobrecarga de ácidos grasos libres, hiperglucemia) pueden ser mecanismos que contribuyen a la aparición de la insulino-resistencia y a otros desórdenes metabólicos (Kim et al 2008). Otros factores genéticos que pueden influir en la contribución del metabolismo mitocondrial a la aparición de la DMT2, son los relacionados con las enzimas de la OXPHOS, y la expresión de la quinasa 1 inducida por PTEN (PINK1), que están suprimidas en músculo de pacientes con DMT2 (Schiff 2009).

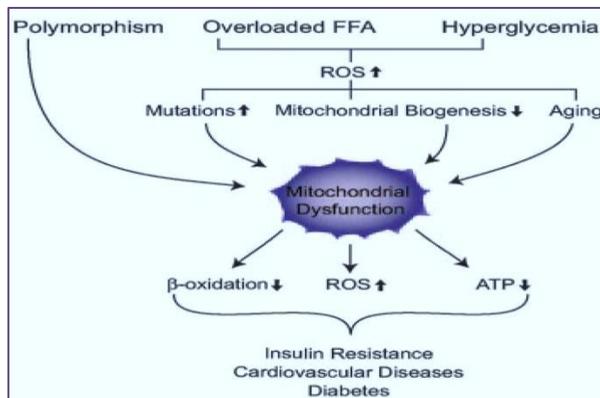


Figura 19. Mecanismo de disfunción mitocondrial. El exceso de nutrientes, aumenta la producción de ROS y reduce la biogénesis mitocondrial, causando alteraciones en la función mitocondrial. Entonces, se produce una disminución en la β -oxidación y en la producción de ATP, que junto con el aumento de ROS producen insulino-resistencia, diabetes y enfermedades cardíacas. 2008 Feb 29; 102(4): 401-414.

Mutaciones del mtADN, como la A3243G, son causales de diabetes en la población general (Smith et al 1999, Maassen et al 1996 y Saker et al 1997). Esta mutación se ha descrito en familias con DMT2 de aparición en adultos y con sordera neurosensorial (Reardon et al y van den Ouweland et al 1992). La patogenia está relacionada con un fallo de la célula beta, con evidencia de reducida secreción de insulina en presencia de una sensibilidad a la insulina normal (Kadowaki 1994). La misma mutación puede resultar en un fenotipo menos común, la encefalopatía mitocondrial con acidosis láctica y episodios similares a infartos (MELAS). Existen otras mutaciones puntuales y deleciones asociadas con diabetes y sordera presentes también en el gen de tRNA de la leucina (Morten et al y Moraes et al 1993, Suzuki 1996), además de las halladas en el gen del tRNA de la lisina (Kameoka et al 1998). El mtADN posee una limitada capacidad de codificar proteínas, por lo que son los genes nucleares los que hacen la mayor contribución a la arquitectura, la biogénesis y los sistemas metabólicos mitocondriales. Varias condiciones relacionadas con la susceptibilidad a la insulino resistencia, tales como la OB, la acumulación lipídica y el envejecimiento, han sido asociadas con una reducción en la expresión de genes de la OXPHOS codificados en el núcleo celular (Crunkhorn 2007 y Sparks 2005). Modificaciones genéticas y epigenéticas podrían estar contribuyendo a la reducción de la expresión de genes de la OXPHOS en la insulino resistencia y la DMT2. Un ejemplo es la expresión del gen COX7A1 del complejo IV, que está reducida en los pacientes con DMT2, y cuya heredabilidad fue estimada en un 50-72% en gemelos (Rönn T 2008). Este hecho indicaría una fuerte contribución genética y ambiental compartida en la familia. Patrones similares se han hallado en los genes NDUF6, perteneciente al complejo I (Ling 2004), y en ATP5O, que codifica para un componente de la ATP sintasa (Rönn 2009). Debido a que se ha visto un incremento de la metilación del ADN de forma paralela a la disminución de la expresión en ciertos genes como COX7A1, NDUF6 y ATP5O, mecanismos epigenéticos podrían también estar involucrados en la reducción de la expresión de genes de la OXPHOS en la insulino resistencia (Ling 2007, Ling y Groop 2009).

A pesar de la importante relación entre la CMR y el metabolismo energético, este último muy afectado en la DMT2, no se han realizado muchos estudios de asociación de genes pertenecientes a este sistema biológico y la DMT2 (Rönn 2009 y Ling 2007). Se han hallado asociaciones entre la DMT2 y polimorfismos de genes nucleares cuyas proteínas regulan la función mitocondrial, como son el coactivador 1 del receptor gama activado por el proliferador del peroxisoma (PPARGC1A) y el factor 1 nuclear respiratorio (NRF1) (Ek et al 2001, Hara et al 2002 y Cho et al 2005), genes que regulan la expresión de los genes mitocondriales. Estudios GWAS recientes, tales como el llamado Diabetes Genetic Initiative (DGI), han contribuido a dar nuevos conocimientos de la naturaleza de los factores genéticos involucrados en la DMT2 (Saxena et al 2007 y Voight et al 2011). En el trabajo de Olsson et al del 2011 se examinaron los resultados del DGI seleccionando sólo los polimorfismos de genes nucleares codificantes para la CMR, para encontrar asociación con éstos y la secreción de insulina. Incluyendo los SNPs con una frecuencia del alelo minoritario (MAF) mayor que 0.05, se ha encontrado asociación entre la secreción de insulina estimulada por glucosa durante la prueba de tolerancia a la glucosa (PTG), representada con el índice insulínogénico, y 9 SNPs de 6 genes de la CMR. Dos de estas variaciones, pertenecientes a los genes COX7A2 (rs1323070) y NDUFC2 (rs606164), se estudiaron en otra población y sus asociaciones se replicaron, confirmando su relación con la secreción de insulina, y su posible influencia en el riesgo de padecer DMT2.

2. Hipótesis

La cadena mitocondrial respiratoria es el principal productor de ATP y de ROS en la célula. Por ello, aporta la energía necesaria para el funcionamiento celular y es un importante regulador del estrés oxidativo.

La DMT2 es una enfermedad metabólica muy compleja, de la cual queda mucho por dilucidar en relación a las alteraciones de complejos biológicos involucrados en su aparición. En esta patología existe una desregulación del metabolismo energético, a nivel mitocondrial, de la producción de ROS y del envejecimiento celular. Además de la diabetes, la hipertensión y la obesidad se han asociado con la disfunción mitocondrial, pero se desconocen las bases moleculares de estos hallazgos.

En base a todo lo anterior, nuestra hipótesis es:

Alteraciones en la cadena respiratoria mitocondrial pueden estar involucradas en el desarrollo de la diabetes mellitus tipo 2 a través de un efecto directo, o al actuar sobre factores de riesgo que influyen en esta enfermedad. Así, variaciones en los genes que codifican las proteínas implicadas en dicha cadena pueden modular su actividad y, por tanto, contribuir al riesgo de desarrollar DMT2.

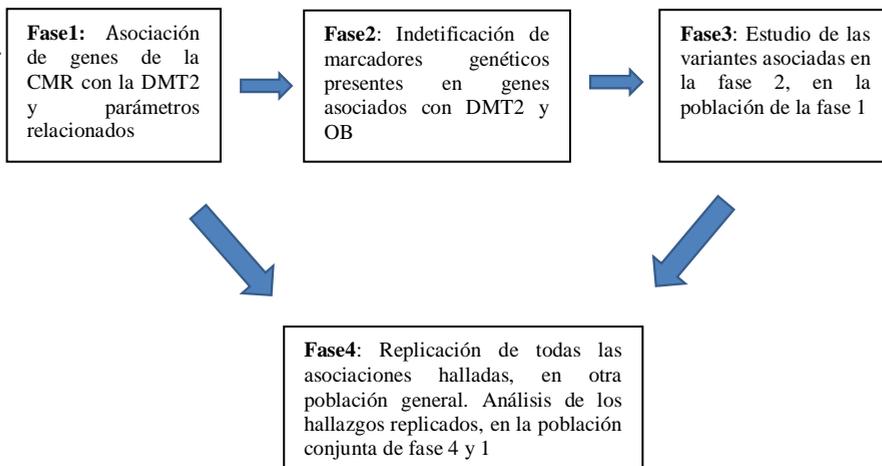
3. Objetivos

- Identificar genes pertenecientes a la cadena mitocondrial respiratoria que muestren una posible asociación con el riesgo de padecer DMT2 y otros desórdenes relacionados, como la obesidad y la hipertensión arterial, en la población española.
- Estudiar la asociación de los polimorfismos de los genes indicados, y parámetros antropométricos y clínicos (índice de masa corporal, circunferencia de la cintura, glucemia).
- Analizar los diferentes marcadores en los parámetros anteriores, en subgrupos de la muestra, separados por sexo y por presencia/ausencia de obesidad.
- Verificar los hallazgos obtenidos en el estudio preliminar, estudiando otra muestra representativa de la población española: el Estudio Di@bet.es.

4. Metodología

4.1 Diseño de la tesis

En este trabajo se incluyeron las tres poblaciones generales descritas a continuación, para poder realizar una primera aproximación al estudio de las asociaciones entre genes de la CMR y la diabetes mellitus tipo 2 a nivel poblacional. Para ello, se realizó el genotipado de distintos polimorfismos de los genes de interés, en la muestra total. Evaluando las asociaciones obtenidas, se seleccionaron genes cuyos SNPs se asociaban con DMT2 y parámetros relacionados, para su estudio posterior más a fondo, mediante la secuenciación de sus promotores y exones. A continuación, se realizó el genotipado de todos los SNPs asociados con DMT2, OB y HTA en otra población general, para replicar en ella las asociaciones halladas previamente. Finalmente, las variantes replicadas se analizaron en la población total de este trabajo, para validar los hallazgos. Un esquema de las distintas fases realizadas en este trabajo se muestra a continuación:



4.2 Descripción de las poblaciones analizadas

Todos los individuos incluidos en las poblaciones analizadas en el presente trabajo dieron su consentimiento para participar en los estudios y los protocolos fueron aprobados por los correspondientes comités éticos.

4.2.1 Población General del Estudio Hortega

La zona en la que se reclutaron los sujetos para el estudio Hortega se localiza en el centro de España. Se seleccionaron individuos al azar de las bases de datos públicas de diferentes registros locales de la provincia de Valladolid, cuya capital posee 214.445 habitantes según el Instituto Nacional de Estadística (datos oficiales del 2005). La población general que se obtuvo incluye a participantes de edades comprendidas entre los 14 a 85 años. Las personas seleccionadas fueron invitadas a participar en el estudio en dos fases: primero se les envió por correo un cuestionario acerca de factores cardiovasculares, y después se seleccionaron al azar 1500 individuos de los que respondieron a ese correo. Si el contacto no se establecía o el individuo se negaba a participar, el investigador podía invitar a un reemplazo individual de una lista de reserva confeccionada al azar. El treinta y dos por ciento de los sujetos contactados inicialmente fueron sustituidos.

El reclutamiento de esta población se llevó a cabo en el Servicio de Medicina Interna del Hospital Universitario del Río Hortega en Valladolid. El trabajo fue dirigido por el doctor Juan Carlos Martín Escudero y en él participaron un importante número de facultativos (Mena-Martin et al, 2003).

Los criterios de exclusión fueron:

- presencia de enfermedades o trastornos graves que pudieran influir en la recopilación de información fiable
- cualquier condición social o mental que pudiera complicar o impedir la participación del sujeto en el estudio

Este estudio se llevó a cabo gracias a la concesión del proyecto “Estudio Hortega” financiado por fondos públicos (Consejería de Sanidad y Bienestar Social de la Junta de Castilla y León, Facultad de Medicina de la Universidad de Valladolid y Hospital Universitario del Río Hortega).

4.2.2 Población General del Estudio Pizarra

El estudio Pizarra se realizó en la población con este mismo nombre perteneciente a la provincia de Málaga, con un total de 6600 habitantes (Instituto Nacional de Estadística, datos oficiales del 2002). 1.226 personas fueron seleccionadas al azar del registro municipal, e invitadas vía correo a acudir al centro de salud local más cercano para realizarse un examen médico. A las personas que no contestaron se les mandó una carta citándolas de nuevo y, por último, fueron contactadas en sus casas si no se tenía contestación. El rango de edades de los individuos reclutados fue de 18 a 65 años (Soriguer-Escofet et al, 2002).

Los criterios de exclusión fueron:

- personas hospitalizadas
- presencia de problemas médicos severos o trastornos psicológicos
- mujeres embarazadas

La participación final fue del 78,5% y la distribución por edad y sexo fue similar a la de la población total de Pizarra (Instituto de Estadística de Andalucía, Sistema de información municipal de Andalucía, Junta de Andalucía).

Este estudio se realizó con las ayudas del Fondo de Investigación Sanitaria (98/0475 and 95/1676), Junta de Andalucía (6/98), del Fondo Europeo de Desarrollo Regional (IFD 97/0718) y de la Asociación Europea Maimónides.

4.2.3 Población General del Estudio Segovia

En este estudio se invitó a participar de manera aleatoria a personas de áreas rurales (poblaciones con 5.000 o menos residentes permanentes) y urbanas de la provincia de Segovia. Según los datos del censo de 2001 (Instituto Nacional de Estadística), la población de Segovia era de 147.694 habitantes, de los que 63.417 tenían entre 35 y 74 años. Una muestra de 2.992 individuos en ese rango de edades se obtuvo a partir de la base de datos proporcionada por el Servicio Regional de Salud mediante sus correspondientes tarjetas sanitarias. Todos los individuos recibieron información personalizada explicando el objetivo del estudio y solicitando la participación voluntaria. En caso de que no hubiera respuesta, se procedió a establecer contacto

telefónico para confirmar o no la participación del individuo invitado. Se reclutaron 1.166 sujetos de los cuales se excluyeron 133 según los siguientes criterios:

- presencia de diabetes tipo 1
- padecimiento de insuficiencia cardíaca o hepática manifiesta
- cirugía en el año previo
- variación del peso superior a 5 kg en los 6 últimos meses
- hospitalización (sujeto ingresado por cualquier razón en un centro hospitalario en el momento de la invitación a participar)

Unos 224 individuos rehusaron entrar en el estudio por diversas razones personales. Por consiguiente, un total de 809 individuos completaron el estudio.

Quince centros de atención primaria de la provincia de Segovia participaron en este estudio (Martínez-Larrad et al, 2005). Este trabajo ha sido subvencionado parcialmente a través de una beca de investigación de Laboratorios E. Lilly Co. España y Bayer AG, España, y por los proyectos FEDER 2FD 1997-2309, FIS 03/1618 y el 03/08 de la Red de Centros RCMN.

Posteriormente, este estudio se amplió hasta llegar a un total de 1241 individuos.

4.2.4 Población General del Estudio Di@bet.es

Entre el 2009 y 2010 se realizó en España, un estudio poblacional a nivel nacional sobre la prevalencia de Diabetes y alteraciones del metabolismo glucídico al que se llamó Di@bet.es (Soriguer et al 2012). Dicho estudio incluyó muestras aleatorias representativas de la población española, en 100 centros de salud. El 55,8% de los pacientes adultos dio su consentimiento para participar. De ellos, un 9,9% fueron excluidos del protocolo siguiendo los siguientes criterios:

- menores de 18 años
- enfermedad severa
- embarazo o parto reciente
- hospitalización

De esta manera, se obtuvieron 5072 individuos adultos, un 41,6% hombres y un 58,4% mujeres. El estudio fue aprobado por el comité ético de investigación clínica

del hospital Carlos Haya, así como de otros comités de ética e investigación clínica regionales de España, obteniéndose un consentimiento por escrito de todos los participantes.

Este estudio fue una iniciativa conjunta de CIBERDEM, la Sociedad Española de Diabetes (SED) y la Federación Española de Diabetes (FED) (actualmente disuelta), con el objetivo de proporcionar un conocimiento exhaustivo de la epidemiología de la diabetes mellitus tipo 2 en España (<http://www.ciberdem.org/programas-de-investigacion/proyectos/estudio-di-betes>).

4.3 Medidas Bioquímicas

De las poblaciones en estudio en este trabajo, a excepción del Estudio Hortega, se obtuvieron muestras de sangre periférica, tras al menos 12 horas de ayuno, para realizar la determinación de valores plasmáticos de parámetros bioquímicos relevantes en los desórdenes metabólicos de interés (glucemia, perfil lipídico, entre otros). Las muestras se procesaron en el laboratorio del centro hospitalario correspondiente, mediante los procedimientos de rutina, sometidos a estrictos controles de calidad.

En el caso del Estudio Hortega, las muestras de sangre se obtuvieron en cualquier momento del día según la disponibilidad de pacientes y del equipo investigador. Tras la realización de los correspondientes análisis, en aquellos pacientes en los que se detectó una alteración en los niveles de glucosa, se realizaba un segundo análisis en ayunas (superior a 12 horas) y sobrecarga oral de glucosa para determinar la presencia o no de diabetes o metabolismo alterado de glucosa.

4.4 Fase 1: Estudio de los polimorfismos de la CMR

4.4.1 Extracción de ADN a partir de sangre periférica

A partir de las muestras de sangre obtenidas, se aislaron las células nucleadas por centrifugación (1000 r.p.m. durante 15 min. a 4 °C) recuperándolas del pellet (buffy

coat). A continuación, se realizó la extracción de ADN de esas células mediante un método comercial estandarizado utilizando el robot Chemagic Magnetic Separador Module1 (Chemagen, Baesweiler, Germany), que permite la extracción automatizada de ADN a partir de 1-10 ml de sangre total o de un pellet de células. El robot utiliza un sistema de partículas magnéticas que permite obtener gran cantidad de ADN y de elevada pureza. Normalmente, a partir de 1 ml de pellet de células se obtienen unos 200 µg de ADN.

4.4.2 Selección de las variantes genéticas

Se realizaron diversas búsquedas bibliográficas en distintas base de datos (Véase Bibliografía) de las asociaciones entre genes de la CMR, la DMT2 y desórdenes relacionados, centrándonos en las relacionadas con aspectos genéticos y genómicos. Se seleccionaron genes que codificaban para distintas subunidades de la CMR, de manera que representaran a varios de los componentes de este sistema codificados a partir de genes cromosómicos. Se elaboró una lista de 29 genes candidatos (Tabla 4) a partir de los cuales se eligieron un total de 96 polimorfismos de un solo nucleótido (SNPs) para ser genotipados en las poblaciones de los Estudios Pizarra, Horteiga y Segovia. En el Anexo I se puede obtener una descripción detallada de los polimorfismos seleccionados.

| HGNC Símbolo | Nombre HGNC | Ensembl ID release 42 | Locus | Ubicación | Número de SNPs por gen |
|--------------|-------------------------------------------------------|-----------------------|------------|---------------------|------------------------|
| BCS1L | BC1 (ubiquinol-cytochrome c reductase) synthesis-like | ENSG00000074582 | 2q33 | 219232623-219236410 | 3 |
| COX8C | cytochrome c oxidase subunit 8C | ENSG00000187581 | 14q32.13 | 92883290-92884453 | 1 |
| COX4I1 | cytochrome c oxidase subunit IV isoform 1 | ENSG00000131143 | 16q22-qter | 84390697-84398109 | 2 |
| COX5A | cytochrome c oxidase subunit Va | ENSG00000178741 | 15q24.1 | 72999672-73017438 | 5 |
| COX5B | cytochrome c oxidase subunit Vb | ENSG00000135940 | 2cen-q13 | 97628953-97631086 | 2 |
| COX6A1 | cytochrome c oxidase subunit VIa polypeptide 1 | ENSG00000111775 | 12q24.2 | 119360287-119362915 | 4 |

| | | | | | |
|--------|-----------------------------------------------------------------------------------|-----------------|------------|---------------------|---|
| COX6A2 | cytochrome c oxidase subunit VIa polypeptide 2 | ENSG00000156885 | 16p | 31346556-31347222 | 1 |
| COX6B1 | cytochrome c oxidase subunit VIb polypeptide 1 | ENSG00000126267 | 19q13.1 | 40830995-40841523 | 1 |
| COX6B2 | cytochrome c oxidase subunit VIb polypeptide 2 (testis) | ENSG00000160471 | 19q13.42 | 60552888-60557994 | 1 |
| COX6C | cytochrome c oxidase subunit VIc | ENSG00000164919 | 8q22-q23 | 100959548-100975071 | 2 |
| COX7A1 | cytochrome c oxidase subunit VIIa polypeptide 1 (muscle) | ENSG00000161281 | 19q13.1 | 41333664-41335611 | 3 |
| COX7A2 | cytochrome c oxidase subunit VIIa polypeptide 2 (liver) | ENSG00000112695 | 6q12 | 76004111-76023212 | 3 |
| COX7B2 | cytochrome c oxidase subunit VIIb2 | ENSG00000170516 | 4p12 | 46431601-46606009 | 1 |
| COX7C | cytochrome c oxidase subunit VIIc | ENSG00000127184 | 5q14 | 85949540-85952337 | 3 |
| GDF9 | growth differentiation factor 9 | ENSG00000164404 | 5q31.1 | 132224772-132230228 | 2 |
| NDUFS1 | NADH dehydrogenase (ubiquinone) Fe-S protein 1, 75kDa (NADH-coenzyme Q reductase) | ENSG00000023228 | 2q33-q34 | 206696080-206732432 | 3 |
| NDUFS2 | NADH dehydrogenase (ubiquinone) Fe-S protein 2, 75kDa (NADH-coenzyme Q reductase) | ENSG00000158864 | 1q23 | 159435729-159450806 | 7 |
| NDUFS3 | NADH dehydrogenase (ubiquinone) Fe-S protein 3, 75kDa (NADH-coenzyme Q reductase) | ENSG00000213619 | 11p11.11 | 47557208-47562689 | 5 |
| NDUFS4 | NADH dehydrogenase (ubiquinone) Fe-S protein 4, 75kDa (NADH-coenzyme Q reductase) | ENSG00000164258 | 5q11.1 | 52892242-53014922 | 5 |
| NDUFS5 | NADH dehydrogenase (ubiquinone) Fe-S protein 5, 75kDa (NADH-coenzyme Q reductase) | ENSG00000164258 | 1p34.2-p33 | 52892242-53014922 | 2 |
| NDUFS6 | NADH dehydrogenase (ubiquinone) Fe-S protein 6, 75kDa (NADH-coenzyme Q reductase) | ENSG00000145494 | 5p15.33 | 1854509-1869161 | 5 |
| NDUFS7 | NADH dehydrogenase (ubiquinone) Fe-S protein 7, 75kDa (NADH-coenzyme Q reductase) | ENSG00000115286 | 19p13.3 | 1334906-1346583 | 6 |
| NDUFS8 | NADH dehydrogenase (ubiquinone) Fe-S protein 8, 75kDa (NADH-coenzyme Q reductase) | ENSG00000110717 | 11q13 | 67554670-67560686 | 4 |

| | | | | | |
|-------|-----------------------------------------------------------------------|-----------------|------------|---------------------|----|
| SDHA | succinate dehydrogenase complex, subunit A, flavoprotein (Fp) | ENSG00000073578 | 5p15 | 218356-256815 | 6 |
| SDHB | succinate dehydrogenase complex, subunit B, iron sulfur (Ip) | ENSG00000117118 | 1p36.1-p35 | 17217817-17253252 | 11 |
| SDHC | succinate dehydrogenase complex, subunit C, integral membrane protein | ENSG00000143252 | 1q23.3 | 159550790-159601154 | 2 |
| SDHD | succinate dehydrogenase complex, subunit D, integral membrane protein | ENSG00000204370 | 11q23 | 111462832-111471727 | 1 |
| UQCRB | ubiquinol-cytochrome c reductase binding protein | ENSG00000156467 | 8q22 | 97311911-97316987 | 3 |
| UQCRQ | ubiquinol-cytochrome c reductase, complex III subunit VII, 9.5kDa | ENSG00000164405 | 5q31.1 | 132230218-132232435 | 2 |

Tabla 4. Genes seleccionados para su estudio en la fase preliminar de este trabajo.

Los SNPs fueron seleccionados en base a los siguientes criterios:

- Distribución en el gen: se seleccionaron los polimorfismos de tal modo que dieran cobertura a todo el gen (desde la región promotora, incluyendo 3000 pb hasta el extremo 3'UTR)
- Posible efecto funcional: aquellos polimorfismos que pudieran tener un efecto funcional más evidente sobre el gen o la proteína y aquellos que se localizaran en regiones reguladoras del gen fueron incluidos (región promotora, 5'UTR, regiones intrónicas a menos de 30 nucleótidos de los exones, exones y 3'UTR)
- Frecuencia del alelo menor (MAF), mayor del 1 % en población caucásica. Datos obtenidos del proyecto Hapmap (The International HapMap Consortium 2003, datos disponibles en el momento del inicio del estudio)
- Según recomendaciones de la casa comercial a la que se le solicitaron los reactivos para el genotipado: la distancia entre SNPs debía ser superior a 20 pb para que la técnica funcione correctamente

Los SNPs seleccionados no cumplían necesariamente todos estos criterios, debido a que, para lograr una buena cobertura a lo largo de cada gen, o para incluir algunos polimorfismos con un posible efecto funcional, se han tenido que seleccionar

polimorfismos con MAFs menores al 1%. Los datos necesarios para la selección, como las secuencias de los genes y ubicación de los polimorfismos, se obtuvieron del National center of biotechnology information (NCBI) GRCh36 y del ensamblaje versión 54.

4.4.3 Método de Genotipado de los polimorfismos de la CMR

Se utilizó la técnica llamada SNPlex (SNPlex™ Genotyping System, Applied Biosystems, Foster) para el genotipado de los 96 SNPs seleccionados (Anexo I) en las 3701 muestras de esta fase. El diseño de las sondas específicas para el estudio de estos polimorfismos lo realizó Applied Biosystems mediante una aplicación específica, ya que vienen incluidas en los reactivos.

Esta técnica utiliza el ensayo de ligación de oligonucleótidos (sondas OLA) múltiple de Applied Biosystems en combinación con tecnología de PCR multiplex, para conseguir la discriminación alélica y amplificar exponencialmente los productos de interés. Esto es posible mediante el uso de un grupo de sondas de ligación específicas de alelo para cada SNP y un set de kits de reactivos universales. En cada reacción es posible determinar 48 SNPs simultáneamente.

Después, para obtener los genotipos, se realiza una electroforesis capilar y un posterior análisis de los resultados en un programa informático, que realiza la detección de los picos electroforéticos, la cuantificación de calidad de esos picos y la asignación de alelos para cada muestra y SNP.

En los pasos de fragmentación del ADN, ligación de las sondas y PCR multiplex, se utilizaron termocicladores GeneAmp PCR System 9700 (Applied Biosystems, Foster).

La electroforesis capilar, se realizó en el equipo ABI3730 (Applied Biosystem, California, EEUU). El software bioinformático utilizado para el análisis de los resultados obtenidos fue el GeneMapper versión 3.7 (Applied Biosystems, Foster).

Los principales pasos de la técnica SNPlex se detallan a continuación.

-Preparación y fragmentación del ADN Genómico

Una vez extraído el ADN genómico, se procedió a realizar los siguientes pasos:

-Cuantificación del ADN midiendo la absorbancia a 260 nm con un espectrofotómetro (Biophotometer, Eppendorf, Hamburg). Se determinó la pureza de la muestra a partir de la relación entre la absorbancia detectada a 260 nm y a 280nm. Posteriormente, las muestras se diluyeron en TE (10mM TrisHCl, pH 8.0 y 1mM EDTA) hasta una concentración final de 100 ng/μl.

-Comprobación de la calidad del ADN genómico: se cargó una alícuota en un gel al 0.8% de agarosa con un 2 % de bromuro de etidio. Si se observaba una banda sólida, se procedía al paso siguiente de fragmentación (figura 20).

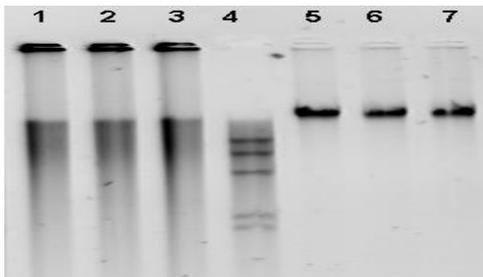


Figura 20. Electroforesis en gel de agarosa. Carrera 4 contiene el marcador de tamaño molecular Λ Hind III. En las carreras 1, 2, 3 se observa ADN fragmentado, mientras que las carreras 5, 6,7 contienen ADN genómico sin sufrir la fragmentación.

- Fragmentación del ADN: Se realizó la fragmentación del ADN genómico por calor para aumentar la eficiencia del genotipado en algunas regiones del genoma, tal como lo indicaba el protocolo. Las condiciones utilizadas en el termociclador fueron las siguientes:

| Pasos | Temperatura (°C) | Tiempo (min) |
|--------------------|------------------|--------------|
| Enfriamiento | 4 | 1 |
| Fragmentación | 99 | 5 |
| Enfriamiento final | 4 | 5 |

Tras la fragmentación, se obtuvo un rango de tamaños de ADN de 2 a 7 kb, que se volvieron a diluir en TE hasta una concentración final de 40 ng/μL.

-Fosforilación y ligación de los oligonucleótidos

En este paso se utilizaron dos oligonucleótidos específicos de los alelos del SNP de interés (OLA), cada uno con el nucleótido de su alelo correspondiente en el extremo 3' (ASO1 y ASO2) para realizar la discriminación alélica de los polimorfismos. Un

tercer oligonucleótido específico de locus, que se une a la región adyacente del SNP (LSO) es necesario, porque contiene el sitio de unión para los cebadores universales de la PCR posterior. Estos tres oligonucleótidos necesitan un paso previo de activación mediante la fosforilación del extremo 5' para ser funcionales. Los linkers universales u oligonucleótidos de enlace son fragmentos de ADN que se unen a una región común a todos los ASO y LSO y evita que éstos se degraden en el paso de la purificación. Información más detallada de estos oligonucleótidos se muestra en la figura 21.

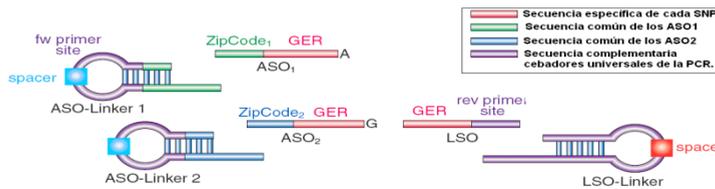


Figura 21. Tipos de oligonucleótidos que componen la mezcla del OLA

La mezcla de reacción por muestra para la fosfoligación es la que figura a continuación:

| Componente de reacción | Volumen por reacción (ul) |
|----------------------------------|---------------------------|
| H2O libre de nucleasas | 0.3 |
| Master Mix de fosfoligación (2X) | 2.5 |
| Linkers universales, 48-plex | 0.05 |
| Sondas de ligación (ASO1-2, LSO) | 0.1 |
| dATP (100X) | 0.05 |
| ADN fragmentado, 40 ng/ul | 2 |
| Volumen Total | 5 |

El OLA consta de una serie de ciclos repetidos, de hibridación de oligonucleótidos y de ligación, para generar múltiples copias de la secuencia de ácidos nucleicos elegida como diana. Los distintos pasos y condiciones son los detallados a continuación:

| Tipo de paso | Temperatura (°C) | Tiempo | Efecto |
|--------------|------------------|--------|---------------------------------------------|
| Incubación | 48 | 30 min | Fosforilación de los oligonucleótidos |
| Incubación | 90 | 20 min | Desnaturalización y activación de la ligasa |
| 25 ciclos | 94 | 15 seg | Desnaturalización |
| | 60 | 30 seg | Hibridación de los oligonucleótidos |
| | 51* | 30 seg | Ligación de los oligonucleótidos |
| Incubación | 99 | 10 min | Inactivación de la ligasa |
| | 4 | | Conservación de las muestras fosfoligadas |

* El paso de Ligación a 51°C se debe producir con una rampa lenta, de un 3%, para que la eficiencia de la reacción sea mayor. Seg: segundos, min: minutos

-Purificación de los Productos de la Ligación

Para purificar el producto de la ligación, se empleó una purificación enzimática combinando dos exonucleasas, para digerir los productos no ligados, oligonucleótidos parcialmente ligados, porciones de los productos de la reacción de ligación que no fueron protegidas por los linkers, y ADN genómico (Figura 21).

La mezcla de reacción de purificación por muestra fue la siguiente:

| Componente de reacción | Volumen por reacción (ul) |
|-----------------------------------|---------------------------|
| H2O libre de nucleasas | 4.2 |
| Buffer SNplex System (10X) | 0.5 |
| Lambda Exonucleasa, SNplex System | 0.2 |
| Exonucleasa I , SNplex System | 0.1 |
| Volumen total | 5 |

Los pasos de la purificación son los que se describen en el siguiente esquema:

5 ul de mix de purificación + 5 ul de producto de ligación



Centrifugación 1 min. a 2000 r.p.m



Incubación en termociclador:

- 90 min a 37°C
- 10 minutos a 80°C (inactivación de las enzimas)
- 4°C hasta proceder al siguiente paso



Dilución 2/5 del producto purificado con H2O libre de nucleasas

-Amplificación de los Productos Purificados de la Reacción de la Ligación

La reacción de PCR contenía las cantidades de reactivos siguientes por muestra:

| Componente de reacción | Volumen por reacción (ul) |
|----------------------------------|---------------------------|
| H2O libre de nucleasas | 2.5 |
| Master Mix de amplificación (2X) | 5 |
| Cebadores de amplificación (20X) | 0.5 |
| Volumen total | 8 |

Se dispensaron 8 ul de la mezcla de reacción en placas de PCR y se agregaron 2 µl del producto de la ligación purificado diluido. Los cebadores para esta reacción son universales, y el cebador 5' esta biotinilado para posibilitar el siguiente paso. Las condiciones de reacción se detallan a continuación:

| Tipo de paso | Temperatura (°C) | Tiempo | Efecto |
|--------------|------------------|--------|------------------------------------------------------|
| Incubación | 95 | 12 min | Desnaturalización y activación de la Taq polimerasa |
| 30 ciclos | 95 | 15 seg | Desnaturalización |
| | 60 | 15 seg | Hibridación |
| | 72 | 45 seg | Elongación |
| Incubación | 72 | 5 min | Extensión final |
| | 4 | | Conservación de las muestras hasta el paso siguiente |

-Captura de los amplicones biotinilados

Se preparó el buffer de unión agregándole por cada 17 ul de este tampón, 0,009 ul del control positivo de hibridación. Se transfirieron 3 µL del producto de PCR biotinilado y 17 ul del buffer de unión a cada pocillo de una placa de hibridación con 384 pocillos recubiertos de estreptavidina (SNplex™ System 384 Well Hybridization Plates, Applied Biosystems, Foster). Así se permitió la captura de los amplicones que tenían el cebador 5' biotinilado, en la placa. La mezcla se incubó durante 30 minutos en constante movimiento rotacional a temperatura ambiente.

Tras la incubación, se añadieron 50 µL de NaOH 0.1 N (Sigma-Aldrich, Steinheim) como agente desnaturizante para separar la doble hélice del ADN y que sólo se

mantuviese unida a la placa una de las hebras. Las hebras que no estaban unidas fueron eliminadas tras varios lavados de las placas de hibridación con 100 μ L una dilución 1:10 de Wash buffer 10X (Applied biosystems, Foster) en agua destilada. El lavado de las placas se realizó en el lavador de placas (ELx405™ Microplate Washer, Bio-tek, Santa Barbara).

-Hibridación de las sondas ZipChute™

En este paso, el producto de la PCR capturado en la placa de hibridación se unió a unas sondas fluorescentes específicas de alelo (Sondas ZipChute), que se utilizaron para decodificar la información de cada polimorfismo estudiado.

Cada sonda ZipChute contiene (Figura 22):

- Una secuencia llamada ZipCode: Esta secuencia es la que se une específicamente al producto de PCR a través de una zona complementaria a una región de ese producto, adyacente al polimorfismo.
- Modificadores de Movilidad: Cada sonda contiene un número diferente de modificadores de movilidad, que permiten su separación por tamaño durante la electroforesis final.
- Un marcador fluorescente: El marcaje fluorescente permite que un instrumento como el ABI 3730 ADN Analyzer, detecte cada sonda.

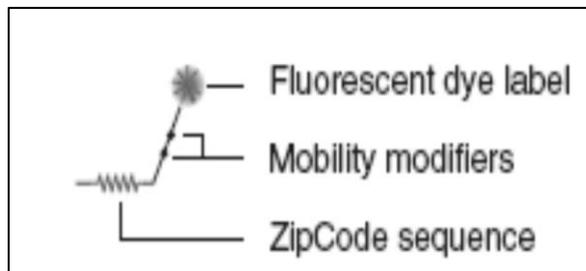


Figura 22. Partes funcionales de una sonda ZipChute.

Las sondas ZipChute están dispuestas en pares, cada una de ellas en representación de cada alelo de un SNP. Cada alelo de un mismo locus puede ser representado por un mismo color fluorescente, por lo que SNPLex usa el tamaño además del color de

las sondas ZipChute para diferenciar los alelos de un locus. Después de la separación por electroforesis capilar en el 3730 ADN Analyzer, se observan unos picos que sirven para identificar los alelos del SNP correspondiente.

Se preparó una mezcla que contenía los siguientes reactivos y cantidades por muestra:

| Componente de reacción | Volumen por reacción (ul) |
|-----------------------------------------------|---------------------------|
| Tampón de dilución de Zipchute, SNPLex System | 13.7 |
| Mezcla de 48 sondas Zipchutes, SNPLex System | 0.05 |
| Desnaturalizante, SNPLex System | 11.25 |
| Total | 25 |

Los pasos que se siguieron en la hibridación fueron los detallados a continuación:

Agregar 25 ul de mix de hibridación a la placa de hibridación



Incubación a 37°C 1 hora en con movimiento rotacional y oscuridad



Lavados con 100 ul del wash buffer 10X diluido 1/100 con agua destilada

La incubación a 37°C durante una hora en constante movimiento rotacional se realizó en un agitador incubador orbital (Shaker Incubator DTS-4, ELMI, Riga).

El lavado de las placas de hibridación para eliminar las sondas no hibridadas se realizó con Wash buffer 10X de Applied biosystems, Foster, en el lavador de placas (ELx405™ Microplate Washer, Bio-tek, Santa Barbara).

-Elución de las sondas ZipChute hibridadas

Se preparó el mix por muestra detallado a continuación para realizar la elución de las sondas ZipChutes presentes en la placa de hibridación:

| Componente de reacción | Volumen por reacción (ul) |
|------------------------------------------|---------------------------|
| Tampón de carga de sondas, SNPLex System | 16.9 |
| Marcador de tamaño, SNPLex System | 0.6 |
| Total | 17.5 |

El marcador de tamaños utilizado es el LIZ® labelled SNPlex™ internal size standard, de Applied Biosystems, Foster city. Con él se realizó una curva de calibración de tamaño para identificar las diferentes sondas Zipchutes presentes.

Se agregaron 17.5 ul del mix de elución a cada pocillo de la placa de hibridación y se realizó una incubación de 15 minutos a 37°C en agitación para separar las sondas de las hebras biotiniladas.

-Electroforesis capilar y análisis de los resultados

Se transfirieron 7 µL de cada pocillo de la placa de hibridación conteniendo las sondas ZipChutes, a una nueva placa para llevar a cabo la separación electroforética en el secuenciador automático ABI3730 ADN Analyser (Applied Biosystems, California, EEUU).

El Sistema de hibridación de SNPlex contiene una mezcla de sondas ZipChute universales. Estas sondas universales constituyen una biblioteca de alelos de referencia, llamadas escalera alélica o marcador de alelos, que se utiliza para normalizar los parámetros de todas las muestras al usar el software GeneMapper versión 3.7 (Applied Biosystems, California, EEUU) para asignar los nombres de cada alelo. Así se permite simplificar esta tarea y realizarla de forma automática. (Fig. 23). En la placa para la separación electroforética se agregaron los marcadores alélicos por duplicado en varios pocillos.

El software GeneMapper® v3.7 se usó para el análisis de los datos obtenidos después de la electroforesis capilar y detección de fluorescencia, llevada a cabo en el 3730 ADN Analyzer en todas las muestras. Este programa bioinformático usa métodos de análisis para la detección de los picos electroforéticos, la asignación de alelos y la cuantificación de calidad de los picos y elabora informes de resultados de calidad. Los picos de cada pareja de locus que se observaron, se utilizaron para identificar los alelos de los correspondientes SNPs.

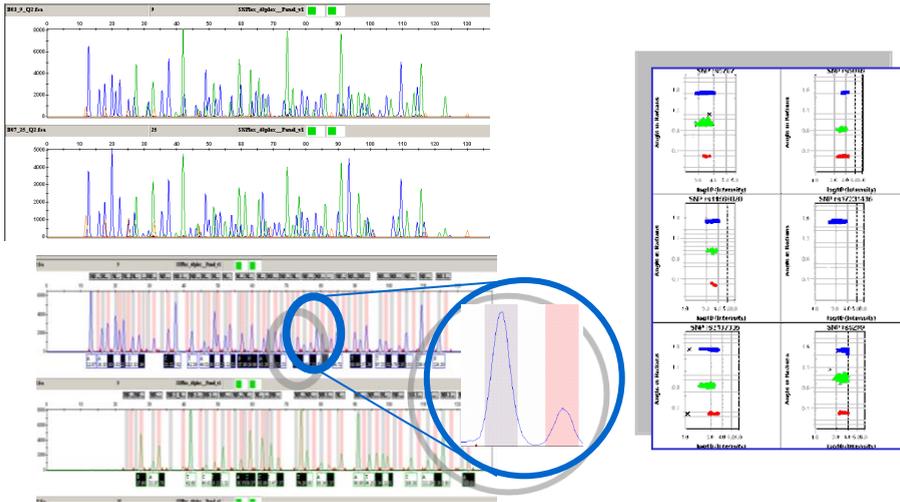


Figura 23. Electroferograma para una separación electroforética de SNIPlex, asignación de alelos para cada polimorfismo y agrupación de genotipos para un polimorfismo concreto.

4.5 Fase 2: Identificación de variaciones genéticas en los genes asociados con DMT2 y OB

Al analizar los resultados de asociación de los SNPs con DMT2 y varios parámetros relacionados, en la población de cribado, se seleccionaron cuatro genes en los cuales se localizaban varios polimorfismos asociados de manera muy significativa con DMT2 y/o OB. Los genes a estudiar más en detalle mediante su secuenciación fueron el SDHB, SDHD, NDUFS5 y NDUFS1.

Se eligieron 100 pacientes dentro de las poblaciones en estudio, formando 2 grupos, uno con sujetos que presentaban DMT2 y/o OB y otro con pacientes sin enfermedad. El criterio diagnóstico para seleccionar a los pacientes diabéticos fue la glucemia en ayunas $>$ a 126 mg/dl o el diagnóstico previo de la enfermedad, y para la OB, un IMC $>$ a 30 kg/m². Estos individuos se seleccionaron según el genotipo que tenían en el SNP asociado significativamente, como se detalla en la tabla 5, con el fin de poder identificar posibles variantes de estos genes asociadas al fenotipo analizado.

| GEN SNP asociado | Nº de individuos | Genotipo | DMT2 | OB |
|----------------------|------------------|--------------|--------------|----|
| NDUFS5 rs11205591 | 50 | GG | NO | NO |
| | 50 | CC | SI | SI |
| SDHD rs10891319 | 50 | AA | NO | NO |
| | 50 | GG | NO | SI |
| SDHB rs1022580 | 50 | 48 GT / 2 GG | NO | NO |
| | 50 | TT | 44 NO (6 SI) | SI |
| NDUFS1 rs3770989 | 50 | TT | NO | NO |
| | 45 | 43 CT / 2 CC | 39 NO (6 SI) | SI |

Tabla 5. Selección de los casos y controles según los genotipos de los polimorfismos asociados con DMT2 y OB, para su posterior secuenciación.

Para la secuenciación de los genes, se siguieron los siguientes pasos:

1. Búsqueda de la secuencia de los genes y localización en ella del promotor, exones y mutaciones presentes (utilización del NCBI GRCh37 y ensembl versión 54).
2. Diseño de los cebadores para la reacción de amplificación (PCR) y de secuenciación de los exones y la región promotora, utilizando el programa informático Primer3 (Rozen y Skaletsky 2000). Cada secuencia obtenida del NCBI, se cargó en el programa primer3, para el diseño de los oligos 5' y 3', indicando las siguientes condiciones de los cebadores:
 - ❖ Temperatura de melting (T_m): entre 60° y 64°, óptima 62°
 - ❖ Diferencia máxima entre T_m de los cebadores 5' y 3': 1
 - ❖ Tamaño del cebador: entre 20y 28 pb, óptimo 24pb

Los demás parámetros no se modificaron respecto a los valores que tiene por defecto el programa Primer3.

3. Después de obtener los cebadores por este programa, se comprobó que éstos eran específicos de la región de interés a amplificar, utilizando el

programa llamado BLAST (Basic Local Alignment Search Tool, <http://blast.ncbi.nlm.nih.gov/Blast.cgi>), que detecta los posibles sitios de unión de los cebadores al genoma, y predice con qué afinidad se unen a ellos. Consideramos que el cebador es lo suficientemente específico para el fragmento deseado, si hay una diferencia en el parámetro E (número de sitios de unión en el genoma, que se esperan encontrar al azar, con el score de alineamiento obtenido para ese fragmento o mayor) de 2 o más órdenes de magnitud, entre él y la siguiente región genómica a la que se une dicho cebador.

4. Puesta a punto de las reacciones de PCR para amplificar la región promotora y todos los exones de los distintos genes. Puesta a punto también de la reacción de secuenciación de todos los fragmentos de interés.
5. Realización de la PCR específica y de la secuenciación de todos los genes en las muestras seleccionadas

Se utilizaron los termocicladores de Applied Biosystems, modelos Veriti y GeneAmpPCRSystem 9700 para placas de 384, para realizar las PCRs y las reacciones de secuenciación. Los reactivos para la PCR de los genes SDHB y SDHD son de Qiagen (Hilden, Alemania), excepto el dNTP que es de Sigma-Aldrich (Steinheim, Alemania). Para los genes NDUFS1 y NDUFS5 todos los reactivos fueron de la marca Euroclone (Milán, Italia) y los mismos dNTPs que para los anteriores.

Para la secuenciación se utilizó el kit BigDye® Terminator v3.1 Cycle Sequencing kit de Applied Biosystems, que incluye el buffer de reacción y el big dye, necesarios para la técnica de secuenciación por terminador marcado con fluorescencia. El Dimetil Sulfóxido (DMSO) es de la casa comercial Sigma-Aldrich. Se realizó esta secuenciación en el equipo ABI3730 (Applied Biosystem, California, EEUU), donde se produce una separación electroforética capilar para la obtención de la secuencia del fragmento a secuenciar.

Los cebadores utilizados en la amplificación de las muestras y en la secuenciación (Sigma-Aldrich, Steinheim) se detallan en la tabla 6.

| Gen | Fragmento amplificado | Secuencias de los cebadores | Tamaño producto PCR |
|--------|-----------------------|--------------------------------------------------------------|---------------------|
| SDHB | Promotor-Exón1 | ccaggaataaagaggtgaaatcg-5' ccctatgcttctcagctctcc-3' | 522 |
| | Exón2 | tatccagcgttacatctgtgtgc-5' ttaagcctccaaggatgtgaaa-3' | 305 |
| | Exón3 | tgacctgagaagaccaaatggata-5' aattaggtgcacagcaagttcac-3' | 336 |
| | Exón4 | ggtgccacatggtcatagagtag-5' atactcaacaatctgccctga-3' | 465 |
| | Exón5 | ggacgagtagtcaggtccaagaa-5' atacactaaatggcttgcacagc-3' | 328 |
| | Exón6 | catgcaagtaggcactttgttcat-5' ctggcttacagcaatctattgtcc-3' | 301 |
| | Exón7 | gatgagcatctaaagctgagtgaa-5' tcaaatgactaggggttctctctg-3' | 347 |
| | Exón8 | gaggaaaggagttcaccacaagat-5' ctgtccatctctccaaggacaag-3' | 394 |
| SDHD | Promotor-Exón1 | agggtgaaactcctttcccttcac-5' gtcattgctgtgatttcggtattt-3' | 489 |
| | Exón2 | gcagatgttccctgttctaact-5' acaggttaggaagctcctaaagg-3' | 386 |
| | Exón3 | ttgggtactgtgtgcatatgtt-5' ttcaacaatttaggcatttca-3' | 339 |
| | Exón4 | ggagacattgcattgaacttgac-5' aggctgtcccaatggactta-3' | 395 |
| NDUFS1 | Promotor1 | ttaacctttccttctcgtggaac-5' tggaggtacgaactaaagacatgc-3' | 595 |
| | Promotor2-Exón1 | ccggaaaattccttatatagagacg-5' cccagtcgaaggacaacagaagac-3' | 655 |
| | Exón2 | aaatcagaacctatgagttcaatg-5' aaaatcatgggttagggctttg-3' | 295 |
| | Exón3 | acaaccataagttgcatccagag-5' tgccacaacaataatgatagctc-3' | 425 |
| | Exón4 | gatgctgtatgtgtgtttcagc-5' gctaggttctccactacgatattgattc-3' | 300 |
| | Exón5 | cctattcttgcctgttatcaaatg-5' ctattgctcctagaagctgagcac-3' | 343 |
| | Exón6- Exón7 | ataggtcattttagtccagcag-5' atcagaacagcatcctcttctc-3' | 515 |

| | | | | |
|--|----------------|---------------------------------------------------------------|---------------------------------------------------------------|-----|
| | Exón8 | tgtttggaaaatcatctcttgg-5' caacagacaaaacaactgaaacc-3' | 384 | |
| | Exón9 | agcataatgagacaccggctctac-5' tcacattataagcaactcagattccag-3' | 579 | |
| | Exón10 | ggtgttagatattgacacttaaccag-5' tgcctcagtagatgtctgaaaatgg-3' | 391 | |
| | Exón11 | agctggggatacaggcgtatg-5' ggatatgttgagaatccaggtg-3' | 549 | |
| | Exón12 | agaaggcaatgggttgatgatag-5' ttgtctccagttatctttggagtc-3' | 350 | |
| | Exón13 | gcctggtctgtagacttctac-5' aacttggttgtatttagtagagacagg-3' | 599 | |
| | Exón14 | ttacacctaaagactgggacaacc-5' catgtgaactgggataatgttc-3' | 391 | |
| | Exón15 | ttgacattgtgaaaggaaattacc-5' aggcaaaattcfaatggaaaac-3' | 365 | |
| | Exón16 | gaagggagttgaattagggttgag-5' aatcatctctgcattcagactgg-3' | 476 | |
| | Exón17- Exón18 | tatttgttcctgcttaactgc-5' cctcaaaagattaagaattgcaaacac-3' | 597 | |
| | Exón19 | taagggtttatcgttggaaatcg-5' aattgaacctctttcatttcac-3' | 597 | |
| | NDUFS5 | Promotor-Exón1 | gggctcatagccatgtgcttaac-5' aacctaaacaaatgtgaaagtggga-3' | 931 |
| | | Seq-promotor | gggctcatagccatgtgcttaac-5' aacctaaacaaatgtgaaagtggga-3' | 640 |
| | | Seq-Exón1 | ctcaacctcctgagtagctggaac-5' aacctaaacaaatgtgaaagtggga-3' | 573 |
| | | Exón2 | aagtttggtagagaagaatcaggag-5' aaagagtcctcaaaaggaacaagaaa-3' | 426 |
| | | Exón3 | ttcttaaaggcggacttcatttc-5' aatctcagtatagaaaggcatgatatgg-3' | 737 |
| | | Seq-Exón3 | ttcttaaaggcggacttcatttc-5' aagaaaagaaagtcgatccaagag-3' | 413 |
| | | Seq-Exón3-cont | tctggagcctgatttctcg-5' aatctcagtatagaaaggcatgatatgg-3' | 404 |

Tabla 6. Cebadores para la amplificación de los genes a secuenciar. Seq: cebadores utilizados solo en la secuenciación.

Los protocolos y cantidades de reactivos utilizados para las reacciones de PCR y de secuenciación, según el gen a amplificar, son los indicados en las tablas 7 y 8 respectivamente.

| GEN | Mix de PCR (por muestra) | Protocolo de PCR | |
|-----------------|------------------------------------|------------------|--------|
| SDHB / SDHD | 1 ul Buffer de reacción 10X | 95°C | 17 min |
| | 0.5 ul Solución Q | 95°C | 15 seg |
| | 0.4 ul MgCl ₂ 25 mM | 62°C | 15 seg |
| | 0.24 ul dNTPs 10 mM | 72°C | 40 seg |
| | 5.11 ul de H ₂ O | 72°C | 3 min |
| | 0.05 ul Taq Polimerasa (Qiagen) | 15°C | ∞ |
| | 1.2 ul de cebadores 3' y 5' (5 uM) | | |
| | 1 ul ADN | | |
| | | } 45 ciclos | |
| NDUFS1 / NDUFS5 | 1 ul Buffer de reacción 10X | 95°C | 3 min |
| | 0.5 ul MgCl ₂ 50 mM | 95°C | 15 seg |
| | 0.2 ul dNTPs 10 mM | 62°C | 15 seg |
| | 5.55 ul de H ₂ O | 72°C | 40 seg |
| | 0.05 ul Taq Polimerasa (Euroclone) | 72°C | 3 min |
| | 1.2 ul de cebadores 3' y 5' (5 uM) | 15°C | ∞ |
| | 1 ul ADN | | |
| | | | |
| | | } 45 ciclos | |

Tabla 7. Protocolo de la PCR realizada en los genes a secuenciar y las cantidades de reactivos y muestra utilizados. Seg: segundos, min: minutos.

| Reactivos de secuenciación | Mix de secuenciación (por muestra y dirección de secuenciación) | | Protocolo de secuenciación | |
|----------------------------|-----------------------------------------------------------------|------|----------------------------|-------------|
| | Buffer de reacción 5X | 1 ul | 96°C | 10 min |
| Big Dye | 0.3 ul | 96°C | 20 seg | } 45 ciclos |
| DMSO | 0.3 ul | 62°C | 2 min | |
| H ₂ O | 0.4 ul | 62°C | 10 min | |
| OLIGOS | 1 ul | 15°C | ∞ | |
| Producto de PCR | 2 ul | | | |

Tabla 8. Protocolo de la secuenciación realizada y las cantidades de reactivos y muestra utilizados. Seg: segundos, min: minutos.

4.6 Fase 3: Estudio de las variaciones identificadas en la fase 2

Después de la búsqueda de los SNPs que mostraron diferencia de frecuencia alélica, entre los pacientes con alguna patología y los controles, se procedió a genotipar esos SNPs en las poblaciones del Estudio Horteiga, Pizarra y Segovia, para estudiar su posible asociación con la DMT2 y desórdenes relacionados.

Los polimorfismos que se seleccionaron a partir de la secuenciación de los 4 genes asociados, para su posterior estudio en la población total de este trabajo, se detallan en la tabla 9.

El genotipado se realizó por secuenciación masiva de los amplicones específicos de cada SNP en el equipo Miseq (Illumina, San Diego, California, EEUU), en la modalidad llamada dual index paired-end. Se utilizó el Miseq Reagent kit v2 de 500 ciclos (Illumina, San Diego, California, EEUU).

Para secuenciar en este tipo de instrumentos, se siguieron estos tres pasos fundamentales:

- 1- Preparación de las librerías a partir de los ADNs de la población total: para obtener las librerías de los 9 amplicones que incluyen en su secuencia los SNPs de interés, se procedió de la siguiente manera:
 - a) Búsqueda de la secuencia y localización genómica de los SNPs de interés (utilización del NCBI y ensembl)
 - b) Diseño de los cebadores para la reacción de amplificación específica (PCR1) a través del programa informático Primer3 (Rozen y Skaletsky 2000). También se utilizó el programa Fast PCR de Java (Kalendar y Schulman 2014) para evaluar la calidad de los cebadores para una reacción multiplex. A estos oligos se les agregó una secuencia universal llamada adaptador, necesaria para el posterior procesado del fragmento obtenido en la PCR1.
 - c) Para la incorporación de un identificador único (índices) a cada muestra, mediante una amplificación (PCR2), se seleccionaron las combinaciones de pares de índices que llevarían cada una de las

muestras de las poblaciones en estudio. Se requiere un par de ellos en cada muestra para realizar la secuenciación en el modo dual index. La secuencia que incluye a los índices, una para el extremo 5' y otra para el extremo 3', lleva una secuencia complementaria para poder unirse a los adaptadores de la PCR1 y con la tecnología de secuenciación de la plataforma Illumina.

| Gen | MAF ncbi | Polimorfismo | Ubicación |
|---------------|----------|----------------------------|----------------|
| <u>SDHB</u> | 0.0429 | rs2746462 (c. 18C>A) | E1 (p.Ala6=) |
| <u>NDUFS5</u> | 0.462 | rs10157617 (c.-482T>C) | 5' del Prom |
| | 0.461 | rs1984600 (c.-140 G>C) | Prom |
| | 0.462 | rs1122472 (c.-3+6G>T) | E1 |
| | 0.015 | rs6981 (c.*18G>A) | 3' UTR |
| <u>NDUFS1</u> | 0.023 | rs77655991 (c.552-24G>A) | intrónico |
| | 0.38 | rs4147718 (c.1026-32G>T) | intrónico |
| | 0.31 | rs1801318 (c.1143 A>G) | E12(p.Arg164=) |
| | 0.38 | rs4147729 (c.1155-101delT) | intrónico |

Tabla 9. Selección de polimorfismos con frecuencias alélicas diferentes entre los casos y los controles del estudio de la fase 2. Prom: promotor del gen, E1: exón 1 del gen.

- d) Puesta a punto de la PCR1 y PCR2 en multiplex (todos los fragmentos de interés se amplifican a la vez para cada muestra en cada PCR).
- e) Realización de la PCR1 en las poblaciones analizadas en este estudio, obteniéndose los amplicones específicos de todos los SNPs a genotipar.
- f) Realización de la PCR2 a partir de la dilución 1/20 de la PCR1.
- g) Agrupación de las PCR2 de las 3701 muestras de la población total tomando una alícuota de cada una (2ul). Se purificó y se cuantificó por fluorescencia.

2- Secuenciación de la muestra

Se cargó el pool obtenido en el secuenciador Miseq, según las instrucciones del protocolo confeccionado por Illumina. Se realizó una carrera llamada dual index paired-end, en donde se utilizaron 2 índices para identificar cada muestra (dual index) y se secuenció a cada amplicón en dirección directa y reversa (paired-end). La secuenciación fue del tipo 2x250, es decir, se secuenciaron 250 bases del amplicón en dirección directa, y 250 bases en dirección reversa.

3- Análisis bioinformático de los resultados

Se analizaron los resultados, obteniendo la cobertura para cada fragmento de cada muestra, para todas las muestras secuenciadas. Las muestras y amplicones con una cobertura mayor a 20x se incluyeron en los análisis de asociación posteriores. Se realizó el análisis bioinformático correspondiente para determinar los genotipos de los SNPs. Se consideraron fiables los genotipos que presentaban un Score de calidad de genotipo (GQ) mayor al 90%.

4.6.1 Preparación de las librerías para la secuenciación masiva

a) Búsqueda de las secuencias genómicas necesarias para genotipar los polimorfismos.

Se utilizó la base de datos del NCBI GRCh37 para buscar la secuencia de los genes en donde se encontraban los polimorfismos, y después, se buscó la ubicación dentro del gen correspondiente, con la ayuda de la base de datos ensembl versión 82. Esto es necesario para poder realizar el diseño de cebadores específicos para la amplificación de las regiones de ADN que incluyan la variación.

b) Diseño de los cebadores para la PCR específica.

Para poder realizar la amplificación de las regiones de ADN en las que se encuentran los 9 polimorfismos en estudio, y para poder secuenciarlos después en el Miseq, se estableció que el tamaño de los amplicones para cada SNP tendría que estar alrededor de 150 pb. La secuencia obtenida del ncbi, se cargó en el programa primer3 para el diseño de los cebadores 5' y 3', indicando las siguientes condiciones de los cebadores:

- ❖ Temperatura de melting (T_m): entre 60° y 64°, óptima 62°
- ❖ Diferencia máxima entre T_m de los cebadores 5' y 3': 1
- ❖ Tamaño del cebador: entre 20y 28 pb, óptimo 24pb

Los demás parámetros no se modificaron respecto a los valores que tiene por defecto el programa Primer3.

Se comprobó que estos cebadores eran específicos de la región genómica de interés a amplificar, utilizando el programa BLAST.

A los cebadores específicos encontrados con el programa Primer3, se les agregó una secuencia adaptador universal para el método Nextera de la plataforma de Illumina, en los extremos 5' y 3', para que puedan ser amplificados e identificados en pasos posteriores, los fragmentos específicos obtenidos con ellos. La secuencia adaptador que se le agregó para cada extremo del fragmento de interés es la siguiente:

Adaptador 5':

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG

Adaptador 3':

GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG

La secuencia completa de los cebadores utilizados en la PCR1, junto con los tamaños de los fragmentos de PCR a obtenerse, se detalla en el anexo II.

Para comprobar que los cebadores no interferían entre sí a la hora de amplificar los 9 fragmentos a la vez, se realizó la prueba *in silico* de los 18 cebadores que se juntarían en la PCR1 en una sola reacción (multiplex), en el programa Fast PCR de Java.

c) Selección de los índices identificativos para cada muestra

Se seleccionaron las combinaciones de pares de índices que llevarían cada una de las muestras de las 3 poblaciones, para su identificación posterior al analizar los resultados de la secuenciación. El índice es un oligonucleótido de 8 pares de bases que se inserta entre dos secuencias llamadas adaptadores. Con el fin de incorporar el índice en la muestra, se realizó una segunda PCR (PCR2) a partir de la PCR1. Para esto, los cebadores de la segunda PCR debían tener en un extremo, la secuencia adaptador que es necesaria para la hibridación a la plataforma de secuenciación, y en otro, la secuencia complementaria al adaptador del cebador de la PCR1, para poder amplificar la secuencia de interés. Las secuencias adaptadores para la incorporación de los índices correspondientes en cada muestra, en la PCR2, son los siguientes:

Secuencia para el extremo 5':

AATGATACGGCGACCACCGAGATCTACAC [i5] TCGTCGGCAGCGTC

Secuencia para el extremo 3':

CAAGCAGAAGACGGCATAACGAGAT [i7] GTCTCGTGGGCTCGG

[i5] hace referencia al índice del extremo 5' y [i7] al índice del extremo 3' de identificación de cada muestra.

Para la secuenciación en la modalidad dual index, cada muestra se identifica con 2 índices, uno en el extremo 3' (i7) y otro en el 5' (i5). Se han utilizado 3701 índices en 5' y la misma cantidad en el 3'.

d) Puesta a punto de la PCR1 y 2

Se realizaron pruebas con distintas concentraciones de cebadores y diferentes condiciones del protocolo de amplificación de la PCR1 y 2, para que se pudieran amplificar los 9 fragmentos de interés al mismo tiempo en una única reacción (PCR multiplex) por muestra. Se utilizó el reactivo para hacer PCRs llamado Qiagen multiplex PCR Kit (Qiagen, Hilden, Alemania). Se comprobó que los 9 amplicones se obtenían en las condiciones testadas realizando una electroforesis multicapilar en gel con el equipo QIAxcel (Qiagen, Hilden, Alemania), donde se produce la separación de los fragmentos según su tamaño, pudiendo diferenciar fragmentos de 3 o más pares de bases de diferencia. Mediante este último método, se comprobó también que no se generaban productos no deseados en las PCRs, como pueden ser los dímeros de cebadores, que pueden obtenerse si no se encuentran las condiciones óptimas para realizar la amplificación.

Las concentraciones y cantidades de reactivos utilizadas para la PCR1 y la PCR2, junto con el protocolo óptimo que se ha obtenido, para la amplificación en los termocicladores, se muestran en la tabla 10.

e) Realización de la PCR1 en la población total

En las amplificaciones de las regiones de interés, en las 3701 muestras de los tres estudios incluidos en la tesis, se utilizaron las siguientes concentraciones y cantidades de los 9 pares de cebadores, para obtener la mezcla de oligos para la amplificación por multiplex:

| Fragmento | Tamaño PCR1 | [inicial] en μM | [final] en μM | Volumen (μl) |
|---------------------------------|-------------|----------------------------|--------------------------|---------------------------|
| NDUFS5_rs10157617 | 185 | 100 | 2 | 200 |
| NDUFS1_rs1801318 | 189 | 100 | 2 | 200 |
| NDUFS1_rs4147729 | 194 | 100 | 2 | 200 |
| NDUFS5_rs6981 | 200 | 100 | 2 | 200 |
| NDUFS5_rs1122472 | 215 | 100 | 2 | 200 |
| NDUFS1_rs77655991 | 218 | 100 | 2 | 200 |
| NDUFS1_rs4147718 | 225 | 100 | 2 | 200 |
| NDUFS5_rs1984600 | 229 | 100 | 2 | 200 |
| SDHB_rs2746462 | 233 | 100 | 2 | 200 |
| H2O | | | | 8200 |
| Volumen total (μl) | | | | 10000 |

Se preparó la mezcla de reacción según el apartado d), para la amplificación de los ADNs de las 3701 muestras totales a estudiar. Se dispensaron 9 μl de mezcla en cada pocillo de placas de 384, y después se le agregó 1 μl del ADN a cada uno.

| | Mix de PCR (por muestra) | Protocolo de PCR | |
|-------------|-------------------------------------------------------------|------------------|----------|
| PCR1 | - 5 μl 2x Qiagen Multiplex Mix | 95°C | 15 min |
| | - 2 μl de H2O | 98°C | 30 seg |
| | -2 μl mezcla de oligos 3' y 5' a 2 μM | 62°C | 30 seg |
| | -1 μl ADN | 72°C | 1 min |
| | | 72°C | 10 min |
| | | 15°C | ∞ |
| | | } 35 ciclos | |
| PCR2 | - 5 μl 2x Qiagen Multiplex Mix | 95°C | 15 min |
| | - 2 μl de H2O | 95°C | 30 seg |
| | -2 μl mezcla de oligos 3' y 5' a 5 μM | 62°C | 1 min |
| | -1 μl de la dilución 1/20 de la PCR1 | 72°C | 1 min |
| | | 72°C | 3 min |
| | | 15°C | ∞ |
| | | } 30 ciclos | |

Tabla 10. Protocolo y cantidades de reactivos y muestra utilizados en la PCR1.

Las secuencias de cebadores utilizadas en la realización de la PCR1 en multiplex y los tamaños esperados para dicha PCR, se muestran en el anexo II.

Se amplificaron las muestras en termocicladores, según las condiciones descritas en el apartado d).

f) Realización de la PCR2 a partir de la PCR1

Se realizó la mezcla de los reactivos necesarios para la amplificación de la PCR2, de las 3701 muestras a estudiar, como se indicó anteriormente en el apartado d, tabla 10. Se dispensaron 7 μ l en cada pocillo de placas de 384, y se les agregó 2 μ l del par de cebadores correspondiente para cada muestra, y finalmente 1 μ l del producto de PCR1, previamente diluido 1/20.

Se amplificaron las muestras en los termocicladores, según las condiciones descritas en el apartado d.

g) Preparación de un único pool de todas las muestras a partir de la PCR2

Se dispensaron 2 μ l de la PCR2 de cada una de las 3701 muestras incluidas en este estudio, para obtener el pool a secuenciar en el Miseq. Se purificaron 3 alícuotas de 20 μ l con el producto de purificación de PCRs llamado MagSi-ADN clean ^{FIX} (MagnaMedics, Geleen, Holanda), en el que el ADN de tamaños mayores a 80 pb se une a la superficie de bolas magnéticas, dejando en solución nucleótidos no incorporados, restos de cebadores, dímeros de cebadores y otros posibles productos no deseados originados en las PCRs. El producto de la PCR2 unido a las bolas se eluyó en el paso final.

Tras la purificación del pool de la PCR2, se realizó la cuantificación por fluorescencia del producto obtenido, utilizando el kit Quantifluor® dsADN system y el sistema de detección “GloMax®-Multi Detection System” (Promega, Madison, EEUU) según las instrucciones de los fabricantes. Este método detecta solamente ADN de doble cadena, al intercalarse el fluoróforo usado para cuantificar, entre las hebras de ADN, lo que da una idea de la cantidad de producto de PCR presente con buena calidad.

Se realizó una dilución del pool a 10 nM, y a partir de ella, otra a 4 nM, como se indica para poder iniciar el protocolo de secuenciación del Miseq.

4.6.2 Secuenciación de las muestras en el sistema Miseq

Se realizó la secuenciación del total de las muestras a partir de la PCR2, en el equipo Miseq, en una carrera llamada dual index paired-end, en donde se utilizaron 2 índices para identificar cada muestra (dual index), y se secuenciaron los amplicones en dirección directa y reversa (paired-end). Se realizó una carrera de 2x250 ciclos, es

decir, se secuenciaron 250 bases en dirección directa (Read1 o Lectura1) y 250 en dirección reversa (Read2 o Lectura2), de la región que incluye el polimorfismo de interés.

Hay dos archivos necesarios para poder iniciar una carrera en un secuenciador de la plataforma Illumina: el documento donde se le asignan a cada muestra, su par de índices identificativos y el método de análisis, así como otros parámetros de la carrera (sample sheet), y el archivo que detalla los nombres y posiciones en el genoma, del inicio y fin de los amplicones a secuenciar (manifest).

A continuación, en la figura 24, se muestra el manifest para la secuenciación realizada en las poblaciones del estudio.

En la figura 25 se muestra el encabezado de la sample sheet, donde están los parámetros que se deben definir antes de cada carrera del Miseq, y se pueden ver las columnas con los nombres de las muestras, y las secuencias de cada índice utilizado para identificarla. Se utilizaron 3701 índices para el extremo 3' y 3701 para el 5', para diferenciar las secuencias obtenidas de cada una de las muestras.

| | A | B | C | D |
|----|---------------------|--------------------------------------------------|----------------|--------------|
| 1 | [Header] | | | |
| 2 | XT Manifest Version | 1 | | |
| 3 | ReferenceGenome | Homo_sapiens\UCSC\hg19\Sequence\WholeGenomeFasta | | |
| 4 | | | | |
| 5 | [Regions] | | | |
| 6 | Name | Chromosome | Amplicon Start | Amplicon End |
| 7 | NDUFS1_rs1801318 | chr2 | 207006768 | 207006889 |
| 8 | NDUFS1_rs4147729 | chr2 | 207006587 | 207006744 |
| 9 | NDUFS5_rs6981 | chr2 | 207003375 | 207003501 |
| 10 | NDUFS5_rs1122472 | chr2 | 207011754 | 207011904 |
| 11 | NDUFS1_rs77655991 | chr1 | 39491523 | 39491640 |
| 12 | NDUFS5_rs10157617 | chr1 | 39492012 | 39492159 |
| 13 | NDUFS1_rs4147718 | chr1 | 39491837 | 39491998 |
| 14 | NDUFS5_rs1984600 | chr1 | 39500114 | 39500246 |
| 15 | SDHB_rs2746462 | chr1 | 17380420 | 17380585 |

Figura 24. Documento con los datos necesarios para iniciar la carrera de secuenciación, y para que, seguidamente, se realice el análisis bioinformático primario.

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|----|----------------|------------------------|-------------|-------------|-------------|----------|-------------|----------|----------|------------|----------------------------------------------|-------------|---|---|
| 1 | [Header] | | | | | | | | | | | | | |
| 2 | EMFileVersion | | 4 | | | | | | | | | | | |
| 3 | Investigator | Pilar-Griselda | | | | | | | | | | | | |
| 4 | Date | 26/07/2016 | | | | | | | | | | | | |
| 5 | Workflow | PCRAmplicon | | | | | | | | | | | | |
| 6 | Application | PCRAmplicon | | | | | | | | | | | | |
| 7 | Assay | NexteraXTv2 | | | | | | | | | | | | |
| 8 | Description | RUN013 | | | | | | | | | | | | |
| 9 | Chemistry | Amplicon | | | | | | | | | | | | |
| 10 | | | | | | | | | | | | | | |
| 11 | [Manifests] | | | | | | | | | | | | | |
| 12 | A | HUMAN.AmpliconManifest | | | | | | | | | | | | |
| 13 | | | | | | | | | | | | | | |
| 14 | [Reads] | | | | | | | | | | | | | |
| 15 | | 151 | | | | | | | | | | | | |
| 16 | | 151 | | | | | | | | | | | | |
| 17 | | | | | | | | | | | | | | |
| 18 | [Settings] | | | | | | | | | | | | | |
| 19 | VariantCalling | Somatic | | | | | | | | | | | | |
| 20 | FlagPCRDupl | 0 | | | | | | | | | | | | |
| 21 | VariantFilter | 30 | | | | | | | | | | | | |
| 22 | outputgenom | FALSE | | | | | | | | | | | | |
| 23 | Adapter | CTGTCTCTATACACATCT | | | | | | | | | | | | |
| 24 | | | | | | | | | | | | | | |
| 25 | [Data] | | | | | | | | | | | | | |
| 26 | Sample_ID | Sample_Nan | Sample_Plat | Sample_Well | Index1 | Index2 | Index3 | Index4 | Manifest | GenomeFolc | Sample_Proj | Description | | |
| 27 | illu-D501 5 | 480001 | 1 | A01 | illu-D501 3 | AGGCTATA | illu-D501 5 | TATAGCCT | A | Homo | sapiens\UCSC\het19\Seouence\WholeGenomeFasta | | | |

Figura 25. Encabezado de la sample sheet utilizada en la carrera de secuenciación

El primer paso en el protocolo de secuenciación del Miseq es la desnaturalización de la librería utilizando NaOH 0.1 N, donde se obtiene como producto final una dilución del pool a 20 pM. Por recomendación de Illumina, para aumentar la heterogeneidad de las librerías, se le agregó un 10 % del fago PHIX (Illumina, San Diego, California, EEUU) a la librería a secuenciar. Para ello, se desnaturalizó de la misma manera que el pool, obteniéndose también una dilución a 20 pM. Este fago es el que utiliza Illumina para poner a punto sus equipos de secuenciación y verificar que éste tenga un correcto funcionamiento.

Para cargar el pool y el PHIX en el secuenciador, se realizó una última dilución de ambos a 8 pM. Después se mezcló la cantidad necesaria de la librería y el fago para tener un 10% de este último en el pool final a secuenciar.

El equipo utiliza la metodología de secuenciación por síntesis (SBS), en la que se secuencia masivamente en paralelo, usando un método reversible basado en terminadores, que permite la detección de cada base de manera individual a medida que se van incorporando a la hebra en síntesis. Esta tecnología del Miseq, utiliza un método de cuatro canales para detectar cada base por separado. Se incorporan los desoxinucleótidos marcados con fluorescencia, utilizando 4 fluoróforos diferentes,

uno para cada tipo de base, lo que conlleva a la adquisición de imágenes en 4 canales de láser distintos por cada ciclo de secuenciación, para la detección del nucleótido que se va incorporando a la hebra que se está secuenciando. El llamado de cada base de la secuencia, se realiza directamente a partir de las medidas de intensidad detectadas en cada ciclo.

A continuación se detalla la técnica de secuenciación de nueva generación SBS.

Para el uso de este tipo de secuenciación de la plataforma Illumina, se necesitan cebadores que incluyan en su secuencia varios oligonucleótidos diferentes en la PCR1 y 2. En la amplificación del segmento de ADN con el polimorfismo (PCR1), se utiliza el cebador específico de la región genómica de interés, al que se le agrega una región llamada adaptador paired-end (PE), para que hibride aquí el cebador de la PCR2. Para la incorporación de los índices al fragmento de la PCR1, los cebadores que se utilizan en la PCR2 tienen, además de las 8 bases que componen el índice, el adaptador PE, y una cola (cola P5 para el extremo 5' y P7 para el 3') que servirá de anclaje a la plataforma donde se secuencia la librería (celda de flujo o flow cell). Los cebadores 3' y 5' de la PCR2 deben tener índices diferentes, y cada muestra tendrá uno de cada clase. La descripción de todas las secuencias que componen una librería para la secuenciación SBS se muestra en la figura 26.

El proceso de secuenciación de la librería final desnaturalizada y diluida, en el secuenciador, consta de los siguientes pasos:

1-Unión de la librería a la celda de flujo mediante la cola de anclaje

La celda de flujo utilizada en la tecnología Illumina tiene inmovilizados sobre su superficie oligonucleótidos directos y reversos. Con ellos se deben hibridar las moléculas de las librerías desnaturalizadas. Esta unión se realiza mediante las regiones del adaptador complementarias a ellos, o sea las colas P5 o P7 descritas anteriormente (figura 27 A).

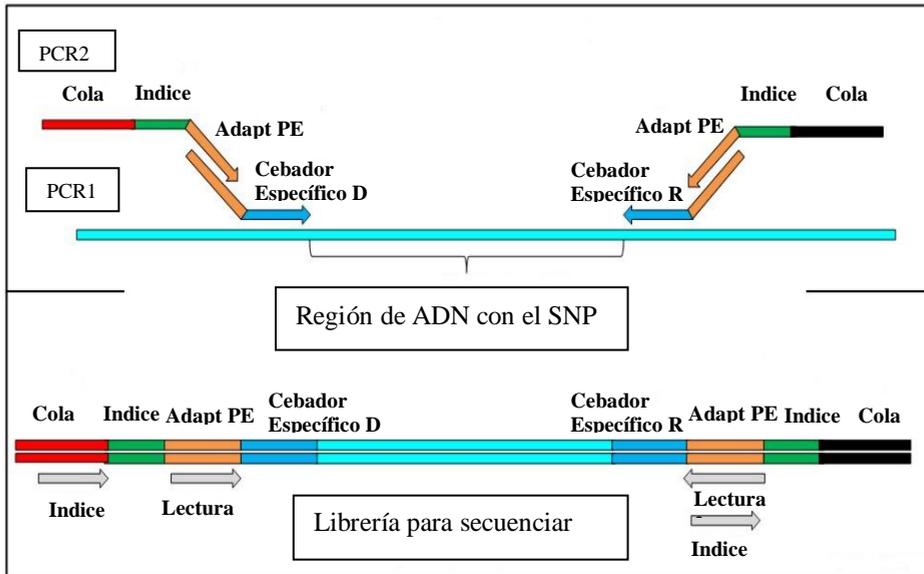


Figura 26. Descripción de las secuencias necesarias para la síntesis de una librería que va a ser secuenciada por una plataforma. (Imagen traducida y modificada a partir de la web Illumina <https://www.fimm.fi/en/services/technology-centre/sequencing/next-generation-sequencing/amplicon-libraries-pcr> (Pekka Elonen FIMM 2013))

2-Amplificación clonal en puente a partir de la hebra unida a la celda de flujo

En este paso, las moléculas unidas a la flow cell se amplifican isotérmicamente mediante la acción de una ADN polimerasa de alta fidelidad, primero desde su extremo libre. Después, al unirse por ese extremo al oligo complementario inmovilizado en la celda de flujo, forman puentes (están unidas las moléculas por su extremo 5' y 3' a la flow cell). Así, se produce una amplificación que genera clones de las distintas librerías, en forma de racimos, llamados clusters. (figura 27 B y C). Se generarán alrededor de 750 millones de clusters, con aproximadamente 1000 moléculas de ADN por cluster (figura 27 D).

3-Linearización de las moléculas amplificadas de ADN, eliminación de la hebra en reverso (reverse) y bloqueo de extremos 3'

Para poder proseguir con la secuenciación, los puentes de clusters deben linearizarse por medio de una desnaturalización, obteniéndose las hebras directa y reversa por separado, unidas a la flow cell (figura 27 E). Como la secuenciación empieza por la hebra directa, la reversa es eliminada (figura 27 F). A continuación, deben

bloquearse los extremos de las hebras hibridadas y de los oligonucleótidos unidos a la celda, para que no se produzca síntesis de ADN allí.

4-Secuenciación de la hebra directa (forward) del inserto de la librería (Lectura1 o Read1)

A la región de ADN que incluye el SNP de interés se la llama inserto de la librería. En este paso, en primer lugar, se produce la unión del cebador de secuenciación, a su secuencia complementaria, ubicada en el adaptador de la hebra inmovilizada, en su extremo libre (figura 28 A). A continuación, tiene lugar la secuenciación por síntesis de las hebras forward, empleando para ello, cuatro desoxinucleótidos trifosfato (A, T, C y G) con grupos terminadores reversibles, marcados con un fluoróforo diferente cada uno. En cada ciclo de secuenciación, los cuatro nucleótidos marcados son incorporados a la superficie a la que están unidos los fragmentos de ADN amplificados. Se produce la unión del nucleótido complementario correspondiente a cada una de las cadenas en formación, y tras su unión, se realiza un lavado para eliminar los nucleótidos sobrantes. A continuación, se recogen las imágenes de fluorescencia para identificar cada base (figura 28 B). Después, se eliminan las moléculas que bloquean la síntesis y que están unidas a los nucleótidos, para permitir la incorporación de la siguiente base en la cadena en síntesis. Este proceso se va repitiendo de forma cíclica, hasta completar 250 ciclos (figura28 C).

5-Secuenciación del índice del extremo 5' ([i5]) de la librería (Index1)

En este paso, se hibrida el cebador para iniciar la secuenciación del índice i5, y a continuación se suceden 8 ciclos de secuenciación, de manera similar a lo detallado para la lectura1, para secuenciar las 8 bases que posee este oligo (figura 28 D).

6-Secuenciación del índice del extremo 3' ([i7]) de la librería (Index2)

Para poder secuenciar el índice i7, la hebra de ADN se hibrida por su extremo libre a la celda de flujo, formando un puente, y se sintetiza una cadena de 8 bases que corresponde a la secuencia del índice del adaptador. Para esto, se realizan 8 ciclos de secuenciación (figura28 E).

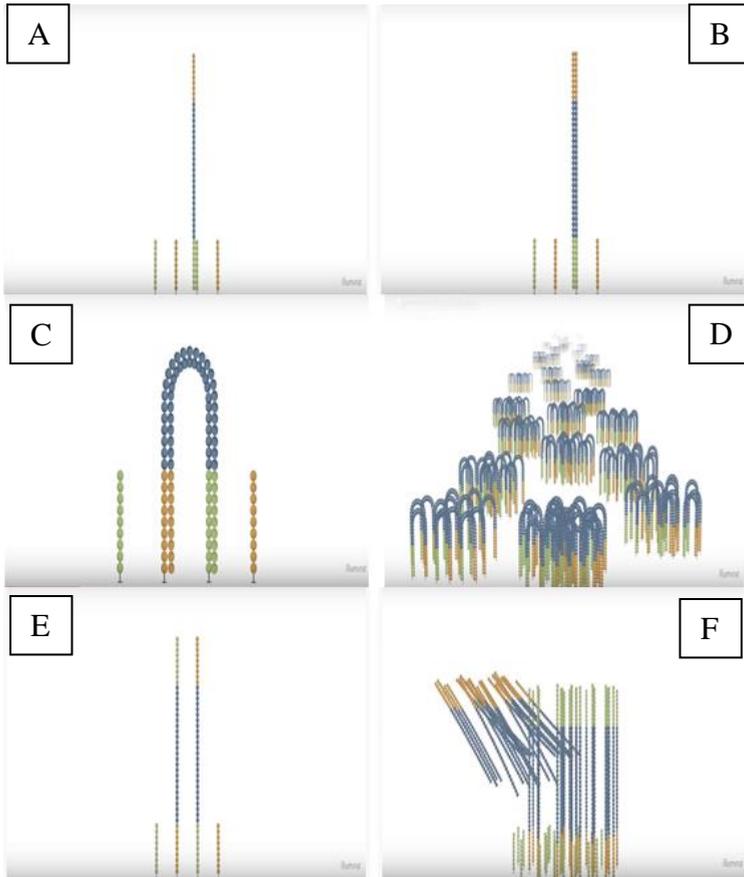


Figura 27. Pasos que componen el proceso de secuenciación de las librerías mediante la plataforma de Illumina, en la modalidad PE-Dual Index. (<https://www.illumina.com/science/education.html>)

7-Amplificación clonal en puente a partir de la hebra reversa (reverse)

En las carreras de secuenciación en las dos direcciones (PE), es necesario obtener la hebra reversa a partir de la hebra forward unida a la celda de flujo. Lo que ocurre aquí es otra amplificación clonal en puente, de la misma manera que en el paso 2, que genera nuevamente los clusters para secuenciar (figura28 F).

8-Linearización de las moléculas amplificadas de ADN, eliminación de la hebra forward y bloqueo de extremos 3´

Los clusters obtenidos en el paso anterior se deben linearizar, es decir, se debe liberar un extremo de la flow cell, y después se procede a la eliminación de la hebra

directa, quedando inmovilizadas en la celda sólo las hebras de ADN reversas. Por último, se bloquean los extremos de las hebras hibridadas y de los oligonucleótidos unidos a la celda, para que no interfieran en la reacción de secuenciación.

9-Secuenciación de la hebra reversa (reverse) del inserto de la librería (lectura2 o Read2)

Primero se une el cebador de secuenciación correspondiente a la lectura2, a la hebra reversa unida a la celda de flujo. Después, se secuencian 250 bases del amplicón, es decir, se realizan 250 ciclos de secuenciación, incorporando en cada ciclo los cuatro nucleótidos con terminadores marcados con fluorescencia, como al realizar la lectura1. Al final de cada ciclo, se recogen imágenes mediante láseres, para detectar qué fluoróforo está presente en los distintos clusters, por ciclo (figura 28 G y H).

4.6.3 Análisis bioinformático de los resultados

El análisis bioinformático de los resultados obtenidos en una secuenciación de nueva generación, se puede dividir en primario, secundario y terciario. Un esquema general de los procesos básicos seguidos en el análisis (primario y secundario), junto con las diferentes herramientas bioinformáticas utilizadas en cada uno de los pasos se representa en la Figura 29.

4.6.3.1 Análisis primario de los resultados de la secuenciación de nueva generación

-Control de calidad de la carrera en el secuenciador

El primer paso para poder obtener los resultados de genotipado, es la comprobación de que la carrera de secuenciación ha funcionado correctamente. La plataforma de secuenciación de Illumina, permite seguir el proceso de secuenciación a tiempo real utilizando directamente el programa con el que se inicia la secuenciación, llamado MiSeq Control Software. Hay varios parámetros disponibles para evaluar la calidad de la carrera. A continuación se detallan los más relevantes.

-Densidad de clusters: número total de clusters que detecta en la celda de flujo el MiSeq, al tomar las imágenes de cada ciclo. Se define como número de clusters por mm^2

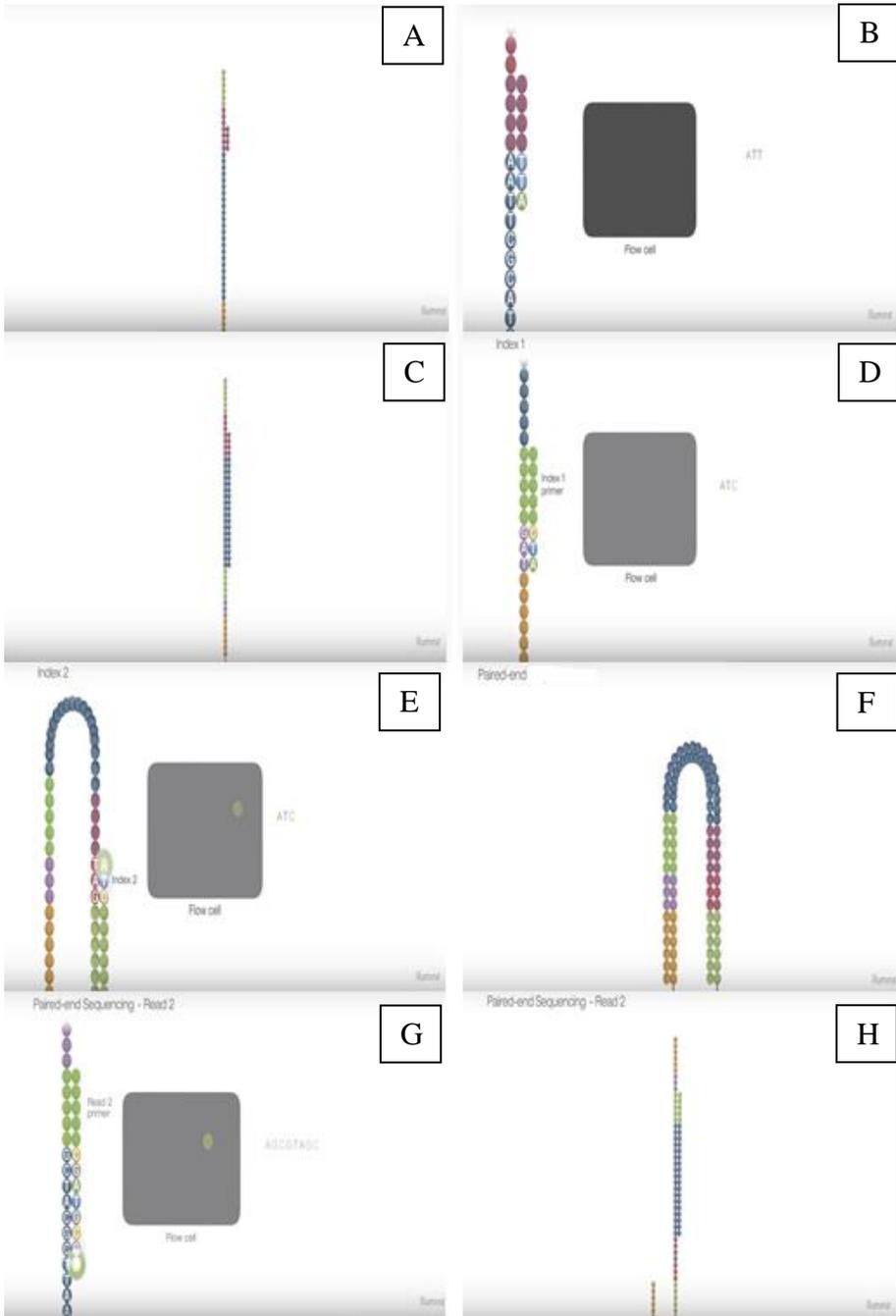


Figura 28. Proceso completo de la secuenciación de la librería en la plataforma Illumina (<https://www.illumina.com/science/education.html>).

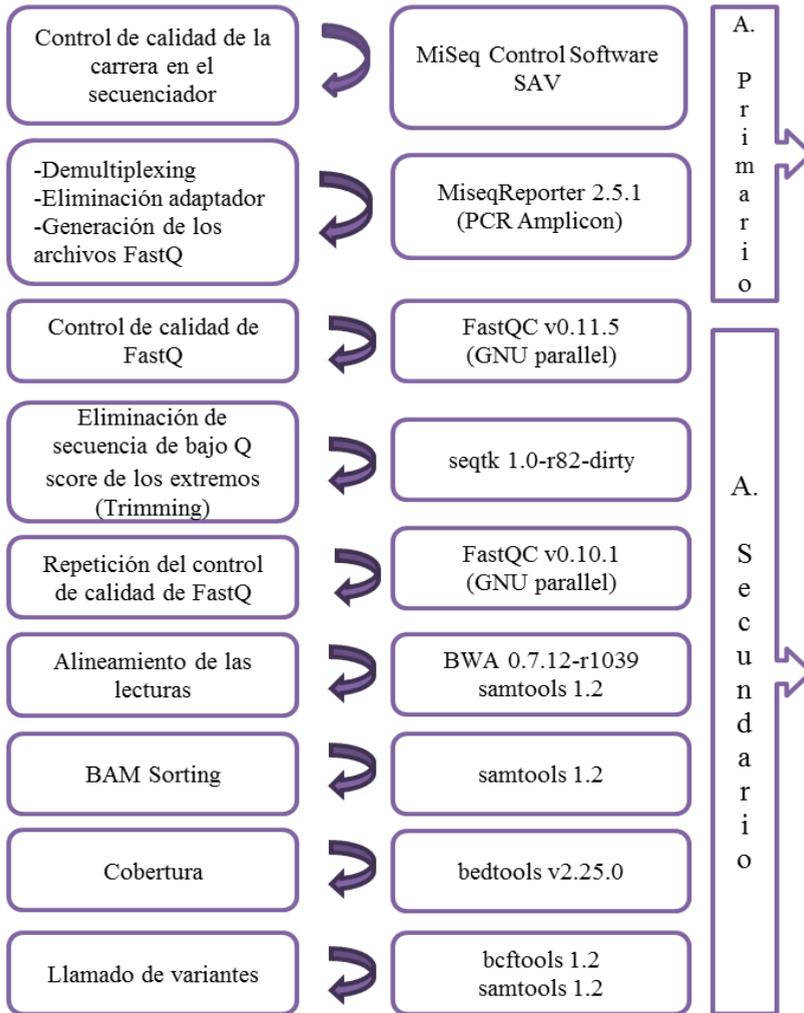


Figura 29. Protocolo (pipeline) del análisis bioinformático de los resultados de la secuenciación.

-Densidad de clusters que pasan el filtro (Clusters PF): número de clusters que cumplen las condiciones de calidad impuestas por Illumina, para considerar fiables los resultados obtenidos a partir de ellos. Se evalúan por mm^2

-Intensidad: señal de emisión que detecta el equipo, proveniente de los fluoróforos de los nucleótidos incorporados en la hebra en síntesis, para los distintos ciclos. En general, debe estar por arriba de 100

-Parámetro de calidad de Phred o Phred quality score (Q score): indica la probabilidad de que una base dada sea nombrada erróneamente por el secuenciador, sirve para evaluar la exactitud de una plataforma de secuenciación. El Q score se obtiene mediante la siguiente ecuación:

$$Q = -10 \log_{10} P \text{ (P es la probabilidad de error en el llamado de una base)}$$

En la tabla 11 se muestra qué exactitud y qué probabilidad de error corresponde a cada Q score. En Illumina se recomienda el uso de los resultados obtenidos por encima de Q30.

| Phred Q Score | Probabilidad de nombrar una base incorrectamente | Exactitud del llamado de bases |
|---------------|--------------------------------------------------|--------------------------------|
| 10 | 1/10 | 90% |
| 20 | 1/100 | 99% |
| 30 | 1/1000 | 99.9% |
| 40 | 1/10000 | 99.99% |
| 50 | 1/100000 | 99.999% |

Tabla 11. Probabilidades y exactitudes del Phred quality score.

- Porcentaje de alineamiento: este parámetro viene determinado por el alineamiento que se realiza de la secuenciación, con el genoma del fago PhiX. Su valor tiene que coincidir con el porcentaje de fago añadido a la librería

-Porcentaje de la tasa de error: este error se calcula según el alineamiento del PhiX. Para una carrera en el Miseq, debe ser menor a 2

-Lecturas totales: para cada equipo, y cada modalidad de carrera disponible, así como para cada versión de reactivos de secuenciación, se debe obtener un número de lecturas estándar, establecido por Illumina según los valores obtenidos al secuenciar el fago PHIX

-Lecturas que pasan el filtro (Reads PF): Las lecturas totales obtenidas, son filtradas por parámetros de calidad establecidos por Illumina

-Porcentaje de lecturas identificadas (Reads Identified PF): este valor es calculado a partir del total de pares de índices que son identificados en la secuenciación, al evaluar las lecturas que pasan los filtros de calidad

-Rendimiento total final (Output): número de bases totales obtenidas en la secuenciación

Las especificaciones para los parámetros más importantes, definidas para el modelo Miseq, utilizando los reactivos v2, para una carrera 2x250 en paired-end, son las descritas en la tabla 12.

| Parámetro | Especificación basada en carrera de PHIX |
|-----------------------|------------------------------------------|
| Clusters PF | 864-965 k/ mm2 |
| Reads PF | 24-30 Millones |
| % Reads Identified PF | 85-90 |
| Output | 7.5-8.5 Gigabases |
| Q score | > Q30 en más del 75% de bases |

Tabla 12. Especificaciones indicadas por Illumina para las carreras 2x250 del Miseq

Existe la posibilidad de usar otro programa, el Sequencing Analysis Viewer (SAV), para evaluar la calidad de la secuenciación cuando la carrera ha finalizado. Aquí se pueden revisar los mismos parámetros descritos anteriormente, pero en más detalle (figura 29).

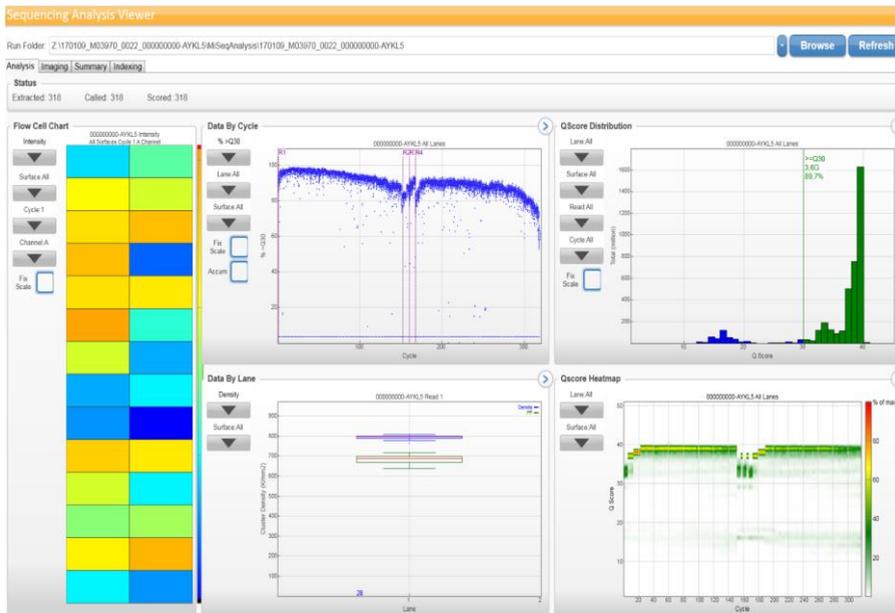


Figura 29. Visualización de los parámetros de la secuenciación mediante el Sequencing Analysis Viewer (SAV).

-Obtención de las secuencias pertenecientes a los amplicones de interés

El programa MiseqReporter (Illumina, San Diego, California, EEUU) se utilizó en primer lugar, para obtener las secuencias de los amplicones de interés, para cada muestra secuenciada. Para esto, el método de análisis llamado PCR Amplicon, con los algoritmos y características definidas por defecto por Illumina, era el adecuado. El programa asigna las lecturas que se obtienen para cada muestra, al reconocer el par de índices que corresponde a cada una de ellas (demultiplexing). A continuación, realiza la eliminación de la secuencia adaptador CTGTCTCTTATACACATCT, que se especificó en la sample sheet, para obtener la secuencia proveniente del amplicón, y no de los adaptadores. Por último, se genera un archivo FastQ para cada muestra, el cual contiene las secuencias obtenidas para esa muestra, y parámetros de calidad de cada lectura. Los documentos FastQ se generan para cada muestra, uno para la lectura 1 y otro para la 2.

4.6.3.2 Análisis secundario de los resultados de la secuenciación de nueva generación

Tras el análisis primario de la carrera de secuenciación, los miembros de la Unidad de Bioinformática del Instituto de Investigación Sanitaria Incliva, procedieron al análisis bioinformático secundario de los resultados obtenidos de la misma.

-Control de calidad de FastQ

Para comprobar la calidad de los documentos FastQ, se utilizó el programa FastQC, versión v0.11.5 (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Se calcularon los parámetros estadísticos de calidad para los FastQ utilizando el siguiente comando genérico:

```
$ fastqc $raw_fastq -o $out_path_preprocessed
```

Para que el comando anterior se ejecutara para todas las muestras secuenciadas, se utilizó el programa GNU parallel (Tange 2011).

El programa calcula los parámetros requeridos para el control de los documentos FastQ para una librería de genoma completo. Un resumen de los parámetros relevantes de calidad a tener en cuenta, para el tipo de secuenciación por amplicones que se realizó, se enumeran a continuación:

- Contenido de bases no identificadas (N) por base secuenciada
- Calidad de secuencia por base
- Scores de calidad por secuencia

-Eliminación de secuencia de bajo Q score de los extremos (Trimming)

En general, los parámetros de calidad disminuyen al acercarnos a los extremos de las lecturas de secuenciación. Por este motivo, se realizó un trimming con el programa seqtk, versión 1.0-r82-dirty (Heng Li, <https://github.com/lh3/seqtk>), para eliminar las bases de ambos extremos, que tuvieran bajos scores de Phred de calidad (valor de corte de la tasa de error de 0.05 y como máximo se cortaron 30 pb). Se utilizó el siguiente comando para realizar el trimming:

```
seqtk trimfq $raw_fastq | gzip > $trimmed_fastq
```

-Repetición del control de calidad de FastQ

Para comprobar si después de eliminar las bases de baja calidad de los extremos, mejoraban los estadísticos de calidad de los FastQ, se repitió el cálculo de dichos parámetros con el siguiente comando genérico:

```
$ fastqc $trimmed_fastq -o $out_path_postprocessed
```

Para que el comando anterior se ejecute para todas las muestras secuenciadas, se utilizó el programa GNU parallel.

-Alineamiento de las lecturas

Para identificar qué lecturas corresponden a cada amplicón secuenciado, deben compararse las secuencias obtenidas, con una secuencia de referencia conocida (mapeo). En este caso, se debe usar el genoma humano de referencia. Se utilizó el programa Burrows-Wheeler Aligner (BWA), versión 0.7.12-r1039 (Li y Durbin R. 2010) y samtools, versión 1.2 (<http://www.htslib.org/>). El mapeo se realizó con la versión de referencia GRCh37 del genoma humano, usando el algoritmo BWA men. De aquí se genera la información para cada muestra en archivos en formato bam. Se utilizó el siguiente comando para realizar el alineamiento:

```
bwa mem -t4 -M -R $RGLINE $REFERENCE_GENOME $trimmed_fastq_R1
$trimmed_fastq_R2 | samtools view -S -b -> $out.bam
```

-BAM Sorting

En este paso se procedió a ordenar los archivos bam obtenidos anteriormente, para poder realizar el llamado de variantes. Se utilizó el programa samtools versión 1.2. El comando genérico utilizado fue el siguiente:

```
samtools sort -T $temp.bam $in_bam -o $out_sorted.bam
```

-Cobertura

Este parámetro es el número medio de lecturas que alinean con (cubren) la secuencia de referencia conocida, y se obtiene a partir de los documentos con formato bam. En nuestro caso, para cada amplicón, tendremos una cobertura distinta, ya que alineará cada amplicón con su región genómica. Por otro lado, cada muestra tendrá su cobertura para cada amplicón. Para su cálculo, se utilizó el programa bedtools versión v2.25.0 (Quinlan A, Hall 2010).

El comando genérico utilizado se detalla a continuación:

```
coverageBed -f $fraction_bed -F $fraction_bam -d -a $bed -b $bam | grep -v '^all\  
> $out
```

-Llamado de variantes

Para el llamado de las variantes estudiadas, se emplearon los programas samtools mpileup (versión 1.2) y BCFtools versión 1.2 (<https://samtools.github.io/bcftools/>). A partir de la información contenida en los archivos bam, se obtuvieron las variantes, para cada muestra, en un archivo denominado vcf (Variant Call Format). El comando genérico utilizado fue el siguiente:

```
samtools mpileup -d 10000 -t DP,DP4 -C50 -E -f REFERENCE_GENOME -u  
$sorted_bam | bcftools call -c -O z -f GQ -o $out.vcf.gz
```

-Filtrado de los genotipos y muestras según su calidad

El parámetro de calidad de genotipo (GQ, genotype quality) se obtiene para cada muestra y SNP en el llamado de variantes. El GQ nos permite predecir si la variación está correctamente nombrada, teniendo en cuenta el score de Phred. El valor de corte utilizado para dar por fiable el genotipo asignado en el análisis, fue un GQ mayor a 90%.

Por otro lado, las muestras para las cuales los amplicones tenían una cobertura que superara el valor de corte de 20 reads (20x), fueron las utilizadas en los pasos posteriores de análisis.

Estos dos valores de corte debían superarse para que los genotipos de los distintos amplicones, en las muestras, se incluyeran en los estudios de asociación.

4.7 Fase 4: Estudios de replicación de las asociaciones halladas

Para poder iniciar la réplica de los SNPs asociados en la fase 1 y 3 con DMT2, OB y HTA, se obtuvieron 4855 ADN's de las muestras sanguíneas obtenidas a partir del Estudio Di@bet.es, para su posterior análisis en este trabajo. Este proceso se realizó en el equipo de extracción de ADN's Maxwell (Promega, Wisconsin, EEUU).

Se seleccionaron 26 polimorfismos que se asociaban muy significativamente con DMT2 u otros parámetros, 23 de la fase 1 y 3 pertenecientes a la fase 3, para proceder a la réplica en el estudio Di@bet.es, de dichas asociaciones. La réplica se realizó mediante el genotipado por secuenciación masiva de los 26 amplicones específicos de cada SNP, en el equipo Miseq (Illumina, San Diego, California, EEUU), en la modalidad llamada dual index paired-end. Se utilizó el Miseq Reagent kit v2 de 300 ciclos (Illumina, San Diego, California, EEUU).

Para secuenciar todas las muestras, se siguieron los mismos pasos que para el genotipado de los SNPs seleccionados, en la fase 3. Un resumen de dichos pasos, y los datos específicos para la réplica, se muestran a continuación:

1-Preparación de las librerías a partir de los ADN's de la población total: para obtener las librerías con los 26 amplicones, que incluyen en su secuencia los SNPs de interés, se realizó lo siguiente:

- a) Búsqueda de la secuencia y localización genómica de los SNPs de interés (utilización del NCBI GRCh37 y ensembl 82)
- b) Diseño de los cebadores para la reacción de amplificación específica (PCR1) a través del programa informático Primer3. También se utilizó el programa Fast PCR de Java para evaluar el comportamiento de estos oligos en una reacción multiplex. Por último, se les agregó una secuencia adaptador universal para el método Nextera de la plataforma de Illumina, para que puedan ser amplificados los fragmentos específicos obtenidos con ellos, en un paso posterior. La secuencia completa de los cebadores utilizados en la PCR1, junto con los tamaños de los fragmentos de PCR a obtenerse, se detalla en el anexo III.
- c) Para la incorporación de un identificador único (índices) a cada muestra, mediante una amplificación (PCR2), se seleccionaron las

combinaciones de pares de índices que llevarían cada una de las muestras de la población en estudio. Esto es necesario para realizar la secuenciación dual index, Esa secuencia se une a adaptadores necesarios para poder hibridar con la PCR1 y con la tecnología de secuenciación de Illumina.

- d) Puesta a punto de la PCR1 y PCR2 en multiplex (todos los fragmentos de interés se amplifican a la vez para cada muestra en cada PCR). En la tabla 13 se detallan las cantidades de reactivos y condiciones óptimas que se encontraron para obtener los 26 amplicones de interés en multiplex.

| | Mix de PCR (por muestra) | Protocolo de PCR | |
|-------------|--------------------------------------|------------------|--------|
| PCR1 | - 5 ul 2x Qiagen Multiplex Mix | 95°C | 15 min |
| | - 2 ul de H2O | 98°C | 30 seg |
| | -2 ul mezcla de oligos 3'y 5'a 5 uM | 62°C | 30 seg |
| | -1 ul ADN | 72°C | 1 min |
| | | 72°C | 10 min |
| | | 15°C | ∞ |
| | | } 18 ciclos | |
| PCR2 | - 5 ul 2x Qiagen Multiplex Mix | 95°C | 15 min |
| | - 2 ul de H2O | 95°C | 30 seg |
| | -2 ul mezcla de oligos 3'y 5'a 5 uM | 62°C | 1 min |
| | -1 ul de la dilución 1/20 de la PCR1 | 72°C | 1 min |
| | | 72°C | 3 min |
| | | 15°C | ∞ |
| | | } 30 ciclos | |

Tabla 13. Protocolos de PCR1 y PCR2 y cantidades de reactivos y muestra utilizadas en la preparación de las librerías.

- e) Realización de la PCR1 en las 4855 muestras analizadas del estudio Di@bet.es, obteniéndose los amplicones específicos de todos los SNPs a genotipar.

Se dispensaron 9 ul de la mezcla de reacción detallada en el apartado d, y 1 ul de ADN en las placas de amplificación.

Se utilizaron las siguientes concentraciones y cantidades de los 26 pares de cebadores, para obtener la mezcla de oligos a utilizar en la amplificación por multiplex:

| Fragmento | Tamaño PCR1 | [inicial] en μM | [final] en μM | Volumen (μl) |
|---------------------------------|-------------|----------------------------|--------------------------|---------------------------|
| rs10891319 | 142 | 100 | 2 | 200 |
| rs39830 | 147 | 100 | 2 | 200 |
| rs4600063 | 154 | 100 | 2 | 200 |
| rs1136224 | 159 | 100 | 2 | 200 |
| rs7255180 | 166 | 100 | 2 | 200 |
| rs13374037 | 172 | 100 | 2 | 200 |
| rs1133322 | 177 | 100 | 2 | 200 |
| rs10742816 | 181 | 100 | 2 | 200 |
| rs17431357 | 184 | 100 | 2 | 200 |
| rs683943 | 187 | 100 | 2 | 200 |
| rs1470625 | 194 | 100 | 2 | 200 |
| rs1053517 | 198 | 100 | 2 | 200 |
| rs4656993 | 202 | 100 | 2 | 200 |
| rs2410718 | 206 | 100 | 2 | 200 |
| rs3770989 | 212 | 100 | 2 | 200 |
| rs535277 | 217 | 100 | 2 | 200 |
| rs9360898 | 221 | 100 | 2 | 200 |
| rs4806187 | 230 | 100 | 4 | 400 |
| rs11205591 | 239 | 100 | 2 | 200 |
| rs17624157 | 242 | 100 | 2 | 200 |
| rs10504961 | 246 | 100 | 2 | 200 |
| rs1022580 | 254 | 100 | 2 | 200 |
| rs803224 | 259 | 100 | 2 | 200 |
| rs77655991 | 218 | 100 | 2 | 200 |
| rs1801318 | 152 | 100 | 2 | 200 |
| rs2746462 | 233 | 100 | 2 | 200 |
| H2O (μl) | | | | 4800 |
| Volumen total (μl) | | | | 10000 |

Antes de utilizar la mezcla de cebadores de la tabla detallada más arriba, en la reacción de amplificación por multiplex del apartado d, ésta se diluyó 1/10 con H₂O.

- f) Realización de la PCR2 a partir de la PCR1 sin dilución.

En este paso, se incorpora en el fragmento de interés obtenido en la PCR1, un único par de índices que identificará a cada muestra por separado, además del adaptador necesario para su posterior secuenciación. Se dispensaron 7 µl en cada pocillo de placas de 384, y se les agregó 2 µl del par de cebadores correspondiente para cada muestra, y finalmente 1 µl del producto de PCR1.

Se amplificaron las muestras en los termocicladores, según las condiciones descritas en el apartado d.

- g) Se agruparon las PCR2 de las 4855 muestras de la población total tomando una alícuota de cada una (2ul). Se purificó y se cuantificó por fluorescencia.

Se dispensaron 2 ul de la PCR2 de cada una de las 4855 muestras incluidas en este estudio, para obtener el pool a secuenciar en el Miseq. Se purificaron 3 alícuotas de 20 ul con el producto de purificación de PCRs llamado MagSi-ADN clean ^{FIX} (MagnaMedics, Geleen, Holanda), de la misma forma que en el apartado 4.6.1.

Tras la purificación del pool de la PCR2, se realizó la cuantificación por fluorescencia del producto obtenido, utilizando el kit Quantifluor® dsADN system y el sistema de detección “GloMax®-Multi Detection System” (Promega, Madison, EEUU).

Se realizó una dilución del pool a 10 nM, y después a 4 nM, según las recomendaciones de Illumina, para iniciar el protocolo de secuenciación del Miseq.

2-Secuenciación de la muestra

Se cargó el pool obtenido en el secuenciador Miseq, según las instrucciones del protocolo confeccionado por Illumina. Se realizó una carrera llamada dual index paired-end, lo que significa que se utilizaron 2 índices para identificar cada muestra (dual index) y que se secuenció a cada amplicón en dirección directa y reversa (paired-end). Se realizó una carrera 2x150 ciclos.

La técnica de la secuenciación por síntesis (SBS) se detalla en el punto 4.6.2, en la fase 3.

3-Análisis bioinformático de los resultados

Se realizó el análisis bioinformático correspondiente para determinar los genotipos de los SNPs de forma similar al realizado en la fase 3. Los pasos que se siguieron están indicados en el apartado 4.6.3.

Para el control de calidad de la carrera en el secuenciador, algunos parámetros que definen si una carrera de 2x150 ciclos es correcta, con los reactivos de Miseq v2, varían con respecto a la de 2x250 realizada en la fase 3. Los valores que cambian para el tipo de carrera utilizados aquí, se detallan en la tabla 14.

| Parámetro | Especificación basada en carrera de PHIX |
|-----------|------------------------------------------|
| Output | 4.5-5.1 Gigabases |
| Q score | > Q30 en más del 80% de bases |

Tabla 14. Valores de Output y Qscore que se deben obtener en la carrera de secuenciación en Miseq para la modalidad 2x150 de PE-Dual Index.

Se consideraron fiables los genotipos que presentaban un score de calidad de genotipo (GQ) mayor al 90%. Las muestras y amplicones con una cobertura mayor a 10x se incluyeron en los análisis de asociación posteriores.

4.8 Análisis estadísticos

Las poblaciones incluidas en este trabajo se analizaron inicialmente de forma individual, y juntando después el total de individuos de los estudios Segovia, Pizarra y Hortega, para utilizar esta población total en el estudio preliminar. Como población para replicar los resultados de asociación obtenidos, se utilizó el Estudio Di@bet.es.

4.8.1 Estadística descriptiva

Se realizó un análisis descriptivo de todos los parámetros relevantes de las poblaciones mediante el programa estadístico SPSS (SPSS International Business Machines, IBM, Nueva York). Se calcularon estadísticos básicos descriptivos como son las medidas de posición (media, mediana), de dispersión (varianza, desviación típica), histogramas y cálculos de intervalos de confianza.

4.8.2 Estudios de asociación

Con los genotipos de los SNPs de la CMR obtenidos por SNPlex, decodificados a partir del programa GeneMapper v3.7, se procedió a realizar el análisis estadístico correspondiente, para asociar la presencia de un genotipo específico, con fenotipos relacionados con la diabetes mellitus tipo 2, en cada población incluida en este trabajo.

Se estimó la prevalencia en la población de estudio para cada alelo y cada posible genotipo, se calcularon las frecuencias alélicas y genotípicas. Antes de empezar el estudio de asociación entre polimorfismos y la DMT2, se realizó el test exacto de equilibrio de Hardy-Weinberg para comprobar la independencia de los alelos heredados por ambos padres. Este test estadístico compara la frecuencia observada del alelo con la esperada asumiendo independencia, con una distribución chi-cuadrado de un grado de libertad. Se tomó como valor de corte, el nivel de significación 0.05 dividido por el número de SNPs estudiados, o sea $0.05/96$ (0.0005) para la fase 1 del análisis, $0.05/9$ (0.0055) para la fase 3, y $0.05/26$ (0.0019) para la fase 4. Se han mostrado en este trabajo, asociaciones con mayor p-valor que el punto de corte descrito, debido a que en un paso posterior se replicarán los resultados obtenidos.

Desde el punto de vista estadístico, un polimorfismo se considera como una variable categórica con un nivel posible para cada genotipo. Una de las categorías se toma de referencia, y suele ser la forma homocigota más frecuente, con la que se comparan el comportamiento de los otros posibles genotipos. Las asociaciones entre polimorfismos y las variables a estudiar, se realizaron utilizando los modelos estadísticos siguientes: si la variable era categórica binaria, se asumió un diseño caso-control y se utilizó el modelo de regresión logística incondicional, si la respuesta era cuantitativa, se consideró una población única y se usaron modelos de regresión lineal. Para las respuestas binarias, se calcularon OR e intervalos de confianza del 95 % además de las frecuencias genotípicas. Para las variables cuantitativas, la regresión lineal se resumió calculando las medias, el error estándar y las diferencias entre las medias respecto a la categoría de referencia, con un intervalo de confianza del 95 %. Se asumió el modelo de herencia codominante en

primer lugar, pero si era posible, los modelos dominante, recesivo y aditivo también se utilizaron, para encontrar el modelo que mejor se ajustara.

Las variables que se analizaron se detallan a continuación:

- Glucemia
- DMT2
- IMC
- Cintura
- OB
- OB Central
- PAD
- PAS
- HTA

Se usaron como covariables la edad y el género en todas las asociaciones estadísticas realizadas. Como el IMC es un factor que influye directamente sobre la DMT2 y la HTA, se usó también como covariable este parámetro, además de la edad y el género, en todas las variables relacionadas con la hipertensión arterial y la diabetes. Esto nos permitió descartar que el efecto del SNP sobre el fenotipo al que se asoció estuviera ligado al IMC. Para el estudio de asociación de los genotipos con la glucemia, PAD y PAS, se eliminaron los pacientes bajo tratamiento: para analizar glucemia se incluyeron los individuos no tratados para la DMT2, y para PAD y PAS los que no recibían tratamiento con antihipertensivos.

Todos los estadísticos antes descritos para el análisis de asociación, se llevaron a cabo con el programa bioinformático SNPStats (Solé et al 2006) del Instituto catalán de oncología (<http://www.gencat.cat>, Cataluña). Este programa utiliza el paquete de R (R Development Core Team, 2005, <http://www.R-project.org>), junto con los paquetes haplo.stats (Sinnwell and Schaid, 2005, <http://mayoresearch.mayo.edu/mayo/research/biostat/schaid.cfm>) y genetics (Warnes and Leisch, 2005). Este análisis se realizó en primer lugar en cada población individual, y posteriormente, en la población total obtenida al juntar los estudios Pizarra, Horteiga y Segovia, para aumentar el tamaño muestral.

A su vez, se realizaron estudios de asociación similares, en subgrupos de la población general total, para hallar asociaciones que afecten a una subpoblación en

particular. Se obtuvieron la población femenina y masculina por separado, para analizar el comportamiento de los SNPs e inferir si el género afecta sus asociaciones. Además, se seleccionaron los pacientes obesos por un lado y no obesos por otro, para buscar los SNPs que se relacionaban con diabetes, en esos subgrupos de la población total, para evaluar una posible interacción entre la asociación y la obesidad.

Para realizar la réplica de las asociaciones halladas en el primer cribado, se realizó el mismo análisis estadístico entre las variables y los SNPs seleccionados por su asociación significativa, en la población Di@bet.es. Los polimorfismos replicados, se analizaron de manera similar en la población conjunta de los 4 estudios incluidos en este trabajo.

5. Resultados

5.1 Fase 1: Estudio de polimorfismos de la CMR

5.1.1 Características generales de las poblaciones analizadas

El análisis de polimorfismos de la primera fase se ha realizado en las muestras de los estudios Horteiga, Pizarra y Segovia. Se han analizado de forma independiente cada población, y posteriormente, se han sumado todas las muestras para comprobar el efecto conjunto. Las características más relevantes de los individuos incluidos en cada estudio, y de todos en su conjunto se muestran en la tabla 15.

5.1.2 Asociaciones de los polimorfismos de la CMR con OB, DMT2 y HTA

De los 96 polimorfismos, pertenecientes a 29 genes de la CMR, que se han seleccionado para el estudio, 7 no han podido ser detectados mediante la técnica SNPlex. Se han excluido 4 SNPs que no cumplían con el equilibrio de Hardy-Weinberg (p -valor > 0.0005 para cumplir este equilibrio). El porcentaje de genotipado ha sido mayor a 90 % en todos los SNPs analizados en las tres poblaciones individuales y en la población total, así como en los subgrupos de la población total que se han estudiado. Unos 23 SNPs no se han incluido en los análisis, debido a que se ha detectado 1 solo genotipo en nuestra población, hecho que podría deberse a una detección incorrecta por esta metodología, de esas variaciones. Finalmente, se han podido utilizar para los análisis de asociación 66 de las variantes.

El estudio de asociación con DMT2, OB y HTA, se realizó primero en cada población individual, con el fin de seleccionar los polimorfismos con asociaciones significativas y con tendencias alélicas similares en los tres estudios. Posteriormente, los SNPs que se comportaban de manera similar en las tres muestras independientes, se analizaron en la población obtenida al juntar los tres estudios. Finalmente, a la población conjunta se la agrupó en población femenina y masculina, y obesa y no obesa por separado, para detectar las asociaciones específicas en estos subgrupos.

| Variables | Población total (n=3731) | Estudio Hortega (n=1502) | Estudio Pizarra (n=988) | Estudio Segovia (n=1241) |
|--------------------------|--------------------------|--------------------------|--------------------------|---------------------------|
| Edad (años) | 51.4 (15.9) | 54.4 (19.3)* | 46.1 (13.9)*** | 51.9 (10.8)** |
| Género (H(%) / M(%)) | 1681(45.1) / 2043(54.9) | 754(50.2) / 748(49.8)* | 365(36.9) / 616(62.3)*** | 562(45.3) / 679(54.7)**** |
| IMC (kg/m ²) | 27.4 (4.6) | 26.4 (4.2)* | 28.6 (5.3)*** | 27.7 (4.3)** |
| Cintura (cm) | 93.6 (13.1) | 89.6 (13.0)* | 99.1 (13.4)*** | 94.3 (11.2) |
| Glucemia (mg/dl) | 95.69 (24.1) | 92.5 (20.4)* | 106.8 (31.7)*** | 91.23 (18.4)** |
| PAS (mm Hg) | 129.6 (20.3) | 130.74 (21.6)* | 132.78 (18.5)*** | 125.7 (19.3)** |
| PAD (mm Hg) | 79.7 (11.0) | 79.1 (10.7)* | 81.7 (11.4)*** | 78.7 (10.8)** |
| OB (N(%)) | 912 (25.3) | 262 (17.4) * | 331 (33.5)*** | 319 (25.7)** |
| OB Centr (N(%)) | 1614 (44.6) | 441 (29.4)* | 599 (60.6)*** | 574 (46.3)** |
| DMT2 (N(%)) | 434 (12.0) | 114 (7.6)* | 194 (19.6)*** | 126 (10.2)**** |
| HTA (N(%)) | 1798 (49.5) | 642 (42.7)* | 607 (61.4) *** | 549 (44.2) |

Tabla 15: Características generales de los estudios Hortega, Pizarra y Segovia individualmente y como una población única. H: Hombre. M: Mujer. IMC: índice de masa corporal. Los valores de esta tabla se expresan como la media (SD) excepto para las variables sexo, OB, OB Central, DMT2, HTA. Diferencias significativas entre las medias se indican de la siguiente manera: * indica $p < 0.001$ entre Hortega y Pizarra, ** $p < 0.001$ entre Hortega y Segovia, *** $p < 0.001$ para Pizarra y Segovia y **** indica $p < 0.05$ comparando Hortega and Segovia.

Los diferentes estudios de asociación se muestran, en ciertos casos, con valores de $p < 0.05$, dado que las asociaciones pueden no ser significativas en todas las poblaciones individuales. Debido a que en un paso posterior se van a replicar los resultados, utilizar estos valores puede permitir detectar polimorfismos con una asociación no muy fuerte en alguna población o en las poblaciones por separado.

5.1.2.1 Asociaciones de los polimorfismos de la CMR y la OB

5.1.2.1.1 Resultados de asociaciones de SNPs de CMR y Obesidad en las poblaciones individuales

Se han obtenido los resultados de asociaciones de los SNPs a estudiar y la OB, en los estudios Pizarra, Hortega y Segovia de forma independiente. En la tabla 16 se

indican los polimorfismos que se asociaban significativamente, o se acercaban al p-valor de 0.05, en alguno de los estudios analizados, y que tenían la misma tendencia en los restantes, al evaluar el riesgo de OB y el IMC.

| Gen SNP | Geno- tipo | ESTUDIO HORTEGA | | | ESTUDIO PIZARRA | | | ESTUDIO SEGOVIA | | |
|-----------------------------|------------------|-----------------|-----------------------------|---------------------------------------|-----------------|---------------------------|----------------------------------------|-----------------|--------------------------|-------------------------|
| | | N | IMC (Kg/m ²) | Obesidad OR/95% CI | N | IMC Kg/m ² | Obesidad OR/95% CI | N | IMC Kg/m ² | Obesidad OR/95% CI |
| SDHC rs4600063 | AA | 1250 | 26.46±0.12 | 1 | 772 | 28.7±0.19 | 1 | 943 | 27.71±0.1 | 1 |
| HWE:1 | AG-GG p-valor | 183 | 26.16±0.3 0.9 | 0.82/0.55-1.22 0.33 | 96 | 27.86±0.4 0.071 | 0.47/0.28-0.78 0.0025 | 132 | 27.21±0.3 0.22 | 0.88/0.57-1.36 0.57 |
| COX7A2 rs683943 | CC-CG | 1410 | 26.4±0.11 | 1 | 865 | 28.58±0.1 | 1 | 1074 | 27.62±0.1 | 1 |
| HWE:0.6 | GG p-valor | 4 | 28.76±2.16 0.3 | 2.65/0.33-21.14 0.36 | 4 | 32.33±0.3 0.044 | NA/0.00-NA 0.001 | 2 | 31.18±4.6 0.22 | 3.17/0.19-52.05 0.43 |
| NDUFS1 rs3770989 | TT | 1308 | 26.41±0.12 | 1 | 799 | 28.49±0.1 | 1 | 1028 | 27.56±0.1 | 1 |
| HWE:1 | CT-CC p-valor | 114 | 26.62±0.44 0.34 | 1.13/0.72-1.77 0.6 | 77 | 29.85±0.6 0.016 | 1.5/0.91-2.48 0.12 | 65 | 27.92±0.6 0.48 | 1.28/0.73-2.23 0.39 |
| NDUFS2 rs1136224 | AA-AG | 1379 | 26.43±0.11 | 1 | 838 | 28.59±0.1 | 1 | 1051 | 27.57±0.1 | 1 |
| HWE:0.36 | GG p-valor | 35 | 26.12±0.79 0.65 | 2.34/1.11-4.91 0.047 | 25 | 29.12±0.7 0.66 | 1.24/0.52-2.95 0.63 | 20 | 28.83±0.9 0.27 | 2.47/1.00-6.06 0.055 |
| NDUFS5 rs11205591 | CC-CG | 1309 | 26.47±0.12 | 1 | 832 | 28.63±0.1 | 1 | 947 | 27.6±0.14 | 1 |
| HWE:0.76 | GG p-valor | 115 | 25.78±0.31 0.049 | 0.67/0.41-1.10 0.11 | 38 | 27.66±0.7 0.38 | 0.86/0.41-1.81 0.69 | 89 | 27.21±0.4 0.25 | 0.89/0.53-1.5 0.66 |

Tabla 16. Asociaciones con el IMC y la OB en las poblaciones individuales de esta fase preliminar.

En el estudio Hortega, se han obtenido asociaciones con la presencia del genotipo GG del SNP rs1136224 y un mayor riesgo de OB, como así también con el genotipo GG del rs11205591 y un menor IMC. Para esos polimorfismos, dichos genotipos han mostrado la misma tendencia en las 2 poblaciones restantes.

En el caso del Estudio Pizarra, el alelo G del polimorfismo rs4600063 se ha asociado con bajo riesgo de OB, mientras que el alelo C del rs3770989 y el genotipo GG del rs683943, han mostrado valores más elevados de IMC. En las 2 poblaciones restantes, estos SNPs han presentado el mismo comportamiento, sin llegar a ser estadísticamente significativo.

Los resultados detallados de las asociaciones de los SNPs con obesidad central y cintura, en las tres poblaciones analizadas, se presentan en la tabla 17.

| Gen SNP | Genotipo | ESTUDIO HORTEGA | | | ESTUDIO PIZARRA | | | ESTUDIO SEGOVIA | | |
|-----------------------------------------|----------|-----------------|--------------|----------------------|-----------------|--------------|----------------------|-----------------|--------------|----------------------|
| | | N | Cintura (cm) | OB Central OR/95% CI | N | Cintura (cm) | OB Central OR/95% CI | N | Cintura (cm) | OB Central OR/95% CI |
| SDHB rs1022580 HWE:0.6 | TT | 1319 | 89.7±0.36 | 1 | 831 | 99.29±0.4 | 1 | 1023 | 89.92±0.3 | 1 |
| | GT-GG | 104 | 88.35±1.36 | 0.86/0.53-1.41 | 49 | 95.88±1.6 | 0.43/0.23-0.81 | 74 | 88.53±1.4 | 0.85/0.51-1.43 |
| | p-valor | | 0.16 | 0.55 | | 0.048 | 0.0091 | | 0.35 | 0.54 |
| SDHC rs4600063 HWE:1 | AA | 1250 | 90±0.37 | 1 | 772 | 99.13±0.4 | 1 | 943 | 89.97±0.3 | 1 |
| | AG-GG | 183 | 87.09±0.95 | 0.65/0.44-0.97 | 96 | 97.75±1.2 | 1.0/0.61-1.62 | 132 | 89.81±1 | 0.85/0.57-1.27 |
| | p-valor | | 0.32 | 0.031 | | 0.16 | 0.99 | | 0.56 | 0.43 |
| SDHD rs10891319 HWE:0.37 | AA | 691 | 88.93±0.5 | 1 | 401 | 98.09±0.6 | 1 | 571 | 89.05±0.5 | 1 |
| | AG-GG | 726 | 90.16±0.48 | 1.04/0.82-1.33 | 478 | 99.97±0.6 | 1.23/0.91-1.68 | 492 | 90.65±0.5 | 1.08/0.83-1.41 |
| | p-valor | | 0.24 | 0.74 | | 0.2 | 0.18 | | 0.028 | 0.57 |
| COX7A2 rs683943 HWE:0.63 | CC-CG | 1410 | 89.56±0.35 | 1 | 865 | 98.97±0.4 | 1 | 1074 | 89.92±0.3 | 1 |
| | GG | 4 | 98.67±2.96 | 6.57/0.37-117 | 4 | 106±3.39 | 6.5/0.54-77.57 | 2 | 93.5±3.52 | 1.68/0.08-33.71 |
| | p-valor | | 0.042 | 0.18 | | 0.094 | 0.12 | | 0.67 | 0.74 |
| NDUFS5 rs11205591 HWE:0.76 | CC-CG | 1309 | 89.76±0.36 | 1 | 832 | 99.01±0.4 | 1 | 947 | 89.84±0.3 | 1 |
| | GG | 115 | 87.89±1.18 | 0.68/0.42-1.09 | 38 | 98.15±1.9 | 1.77/0.79-3.97 | 89 | 88.96±1.2 | 1.14/0.7-1.87 |
| | p-valor | | 0.046 | 0.1 | | 0.91 | 0.16 | | 0.53 | 0.6 |

Tabla 17. Asociaciones con la cintura y la OB Central en las poblaciones individuales de esta fase preliminar.

Se ha encontrado que el alelo G del polimorfismo rs10891319, y el genotipo GG del rs683943, presentaban valores de cintura más altos en las tres poblaciones, mientras que el alelo G del SNP rs1022580, y el genotipo GG del rs11205591, mostraban medidas de cintura más bajas. En la figura 30 se muestra el comportamiento de los valores de cintura, en los distintos genotipos del rs1022580, para las tres poblaciones analizadas. Con un menor riesgo de padecer OB Central, se ha hallado el alelo G de las variantes rs1022580 y rs4600063, siendo estadísticamente significativa dicha asociación en el Estudio Pizarra y en el Hortega, respectivamente.

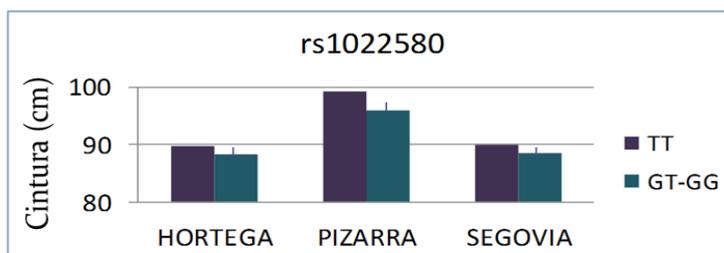


Figura 30. Comportamiento de los valores medios de cintura para los distintos genotipos del polimorfismo asociado, en las tres poblaciones analizadas (SD es cercana a 0 en los genotipos TT y por lo tanto no es visible en el gráfico).

5.1.2.1.2 Resultados de asociaciones de SNPs de CMR y OB en la población total

Una vez obtenidos los resultados de asociación de los estudios de Pizarra, Segovia y Hortega, se han seleccionado todos los SNPs que presentaban alguna asociación y que tenían la misma tendencia en las poblaciones analizadas independientemente, con respecto a su comportamiento con el IMC, cintura, riesgo de OB o de OB Central, para proceder a su estudio en la población conjunta.

En los análisis de asociación de OB e IMC, en la población conjunta de los estudios Pizarra, Hortega y Segovia, se han hallado asociaciones significativas en los 7 SNPs que se seleccionaron al evaluar su comportamiento con dichas variables, a partir de las poblaciones por separado. El rs1136224 no ha mostrado asociación con el IMC al igual que sucedía en las poblaciones independientes. En la figura 31 se muestra el comportamiento de las diferencias entre las medias de IMC, de dos de los SNPs asociados con esta variable, en los estudios analizados por separado y en la población conjunta. Los polimorfismos rs4600063, rs11205591, rs683943 y rs10891319, además de haberse asociado con el riesgo de padecer OB, se han asociado con el IMC. Mientras que en los dos primeros, sus alelos minoritarios mostraban valores bajos de IMC y menor riesgo de OB, en los dos restantes presentaban IMC altos y mayor riesgo de OB.

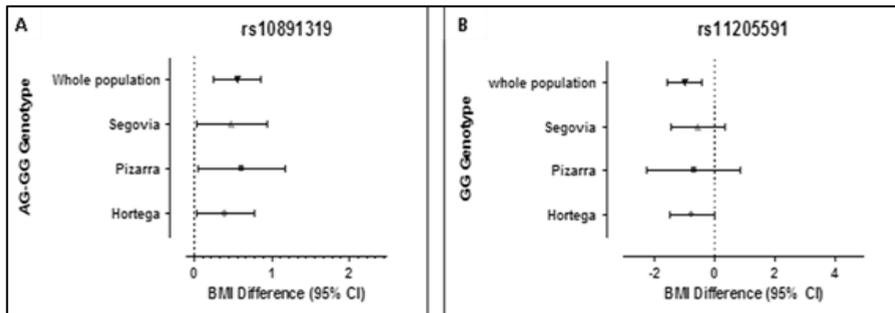


Figura 31. Comportamiento de las diferencias entre las medias del IMC de los genotipos AG-GG del10891319 (A) y del genotipo GG del rs11205591 (B) asociados con esa variable, respecto a los genotipos tomados como referencia, en las poblaciones individuales y en su conjunto. Whole population: población total analizada. BMI: IMC.

De manera similar, hemos hallado asociaciones con Cintura y OB Central en 6 de los 7 polimorfismos seleccionados en la muestra total. Los SNPs rs3770989 y rs1136224 no han mostrado asociación significativa con cintura ni en esta población ni en las poblaciones individuales, al contrario del rs4600063, que al analizar los tres estudios en conjunto, sí lo hizo. Se han relacionado cuatro SNPs con OB Central, sin embargo, sólo los polimorfismos rs1022580 y rs4600063 tenían asociación en alguna de las poblaciones independientes. En la tabla 18 se detallan los resultados obtenidos del análisis de todos los parámetros de obesidad en la población conjunta.

| Gen SNP | Geno tipo | N | IMC (Kg/m2) | Cintura (cm) | No Obesos (n%) | Obesos (n%) | OR/95% CI | No obesidad central (n%) | Obesidad central (n%) | OR/95% CI |
|-----------------------------------------|------------------------|--------------|----------------------------------------|-----------------------------------------|---------------------|----------------------|----------------------------------------------|--------------------------|-----------------------|--------------------------------------------|
| SDHB rs1022580 HWE:0.6 | TT GT-GG p-valor | 3173 227 | 27.4±0.10 26.6±0.28 0.014 | 92.3±0.2 90.0±0.9 0.0062 | 2276/93 170/7 | 892/94.3 54/5.7 | 1 0.82/0.59-1.13 0.21 | 1840/92.6 148/7.4 | 1374/94.5 80/5.5 | 1 0.72/0.53-0.97 0.026 |
| SDHC rs4600063 HWE:1 | AA AG-GG p-valor | 2965 411 | 27.4±0.1 26.8±0.2 0.04 | 92.4±0.2 90.5±0.6 0.014 | 2109/86 318/13.1 | 851/90.4 90/9.6 | 1 0.72/0.56-0.92 0.0072 | 1717/86.7 263/13.3 | 1289/89.6 150/10.4 | 1 0.76/0.61-0.96 0.02 |
| SDHD rs10891319 HWE:0.37 | AA AG-GG p-valor | 1663 1696 | 27.1±0.1 27.6±0.1 0.0004 | 91.2±0.3 93.1±0.3 0.0001 | 1243/51 1178/48 | 422/45.4 508/54.6 | 1 1.25/1.08-1.46 0.0038 | 1008/51.2 962/48.8 | 681/47.6 749/52.4 | 1 1.14/0.99-1.32 0.097 |
| COX7A2 rs683943 HWE:0.63 | CC-CG GG p-valor | 3349 10 | 27.4±0.1 30.7±1.2 0.018 | 92.1±0.2 100.8±2.5 0.044 | 2416/99 3/0.1 | 924/99.2 7/0.8 | 1 6.83/1.72-27.14 0.0039 | 1965/99.8 3/0.2 | 1426/99.6 6/0.4 | 1 4.75/1.04-21.7 0.034 |
| NDUFS1 rs3770989 HWE:1 | TT CT-CC p-valor | 3135 256 | 27.3±0.1 27.9±0.32 0.026 | 92.1±0.2 92.7±0.9 0.29 | 2273/92 173/7.1 | 857/91.2 83/8.8 | 1 1.30/0.99-1.72 0.066 | 1830/92.2 154/7.8 | 1347/92.9 103/7.1 | 1 0.92/0.70-1.21 0.56 |
| NDUFS2 rs1136224 HWE:0.36 | AA-AG GG p-valor | 3268 80 | 27.3±0.1 27.7±0.5 0.52 | 92.06±0.2 94.65±1.6 0.13 | 2364/98 44/1.8 | 897/96.3 34/3.6 | 1 2.02/1.27-3.21 0.0034 | 1921/98.2 35/1.8 | 1390/97.1 42/2.9 | 1 1.78/1.09-2.90 0.02 |
| NDUFS5 rs11205591 HWE:0.76 | CC-CG GG p-valor | 3088 242 | 27.4±0.1 26.6±0.2 0.0011 | 92.28±0.2 89.93±0.8 0.0012 | 2219/92 184/7.7 | 863/94 55/6 | 1 0.72/0.52-0.99 0.039 | 1813/92.5 146/7.5 | 1315/93.1 97/6.9 | 1 0.83/0.62-1.11 0.22 |

Tabla 18. Resultados de las asociaciones del análisis de OB en la población total de la fase preliminar.

5.1.2.1.3 Resultados de asociaciones de SNPs de CMR y OB en subpoblaciones

Hemos procedido a estudiar las variaciones de CMR en subgrupos de la muestra, con el fin de detectar un mayor efecto de alguna de las variables, sobre una subpoblación en particular. En el caso de OB, se analizó a las poblaciones masculina y femenina por separado. En la población diabética, el estudio no lo hemos realizado al haber pocos pacientes con la enfermedad, lo que suponía un bajo poder estadístico si se dividía la muestra.

5.1.2.1.3.1 Resultados de las asociaciones según género

Se analizaron las asociaciones en la población masculina y femenina de la población total, de forma independiente. Los resultados para cada género por separado, para las variables OB e IMC se muestran en detalle en la tabla 19, y para Cintura y OB Central en la tabla 20. Al analizar el IMC, los SNPs rs1022580 y rs3770989 han presentado una asociación más significativa en la población de mujeres, que en la población general. Por otro lado, el rs1136224 muestra asociación en la población masculina, mientras que en la general no la presentaba. Al estudiar el riesgo de OB, hemos encontrado una asociación del alelo C del rs3770989, con un mayor riesgo en las mujeres, a pesar de que en la población general no se viera ninguna.

| SNP | Genotipo | N° de pacientes | | IMC (Kg/m ²) | | No Obesos (N%) | | Obesos (N%) | | OR(95% CI) | |
|-----------|----------|-----------------|------|--------------------------|---------------|----------------|-----------|-------------|------------|-----------------------|-----------------------|
| | | H | M | H | M | H | M | H | M | H | M |
| rs3770989 | TT | 1416 | 1717 | 27.46±0.1 | 27.21±0.1 | 1069 (92) | 1202 (93) | 364 (93.8) | 493 (89.3) | 1 | 1 |
| | CT-CC | 115 | 141 | 27.04±0.3 | 28.64±0.5 | 91 (7.8) | 82 (6.4) | 24 (6.2) | 59 (10.7) | 0.79/0.50-1.26 | 1.89/1.30-2.73 |
| | p-valor | | | 0.31 | 0.0004 | | | | | 0.32 | 0.0009 |
| rs1022580 | TT | 1427 | 1744 | 27.44±0.1 | 27.41±0.1 | 1082 (93) | 1192 (92) | 361 (93) | 531 (95.2) | 1 | 1 |
| | GT-GG | 105 | 122 | 27.23±0.3 | 26.14±0.4 | 78 (6.7) | 92 (7.2) | 27 (7) | 27 (4.8) | 1.04/0.66-1.64 | 0.65/0.41-1.03 |
| | p-valor | | | 0.57 | 0.0054 | | | | | 0.86 | 0.061 |
| rs1136224 | AA | 1142 | 1372 | 27.31±0.1 | 27.34±0.1 | 883 (77) | 949 (74) | 271 (70.4) | 409 (74.9) | 1 | 1 |
| | AG-GG | 369 | 464 | 27.77±0.2 | 27.22±0.2 | 257 (22) | 318 (25) | 114 (29.6) | 137 (25.1) | 1.46/1.12-1.89 | 1.05/0.82-1.34 |
| | p-valor | | | 0.034 | 0.95 | | | | | 0.0052 | 0.7 |

Tabla 19. Asociaciones con IMC y OB en hombres (H) y mujeres (M) de la población total estudiada en la fase preliminar. N: número de muestras.

Con respecto a los valores de cintura, se ha encontrado asociación del rs3770989 en mujeres, que no está presente en el conjunto de la población. Por el contrario, el SNP rs4600063 se asoció en la población total, pero no en el análisis de los géneros por separado. El comportamiento de los genotipos del rs1022580 fue el mismo en mujeres y en población general, con una asociación más significativa en la segunda. En OB Central, el alelo G del rs4600063, se relacionó a menor riesgo, de forma más significativa en mujeres que en la población conjunta de los tres estudios. De manera opuesta, el rs1022580 se asoció con menor fuerza en mujeres que en la población total.

| Población total | Genoti- pos | N° de pacientes | | Cintura (cm) | | No Ob Central (N(%)) | | Ob Central (n(%)) | | OR (95% CI) | |
|-----------------|---------------------|--------------------|------|---------------|-----------------|-------------------------|---------------|-------------------|---------------|-----------------|-----------------|
| | | H | M | H | M | H | M | H | M | H | M |
| rs4600063 | AA | 1360 | 1608 | 96.7 (0.2) | 88.71 (0.36) | 973 (88.3) | 743 (84.7) | 397 (88.2) | 891 (90.2) | 1 | 1 |
| | AG-GG | 182 | 227 | 95.6 (0.7) | 86.33 (0.92) | 129 (11.7) | 134 (15.3) | 53 (11.8) | 97 (9.8) | 1.05(0.74-1.49) | 0.61(0.45-0.82) |
| | HWE:0.64 p-valor | | | 0.26 | 0.063 | | | | | 0.77 | 0.0011 |
| rs11205591 | CC-CG | 1415 | 1675 | 96.6 (0.2) | 88.62 (0.35) | 1007 (92.7) | 806 (92.3) | 416 (94.5) | 899 (92.6) | 1 | 1 |
| | GG | 102 | 139 | 95.2 (0.9) | 86.02 (1.13) | 79 (7.3) | 67 (7.7) | 24 (5.5) | 72 (7.4) | 0.72(0.44-1.16) | 0.91(0.62-1.31) |
| | HWE:1 p-valor | | | 0.18 | 0.011 | | | | | 0.17 | 0.60 |
| rs3770989 | TT | 1432 | 1709 | 96.7 (0.2) | 88.18 (0.34) | 1018 (92.4) | 811 (92) | 421 (93.3) | 925 (92.7) | 1 | 1 |
| | CT-CC | 112 | 140 | 95.0 (1.0) | 90.76 (1.32) | 84 (7.6) | 70 (8) | 30 (6.7) | 73 (7.3) | 0.91(0.58-1.42) | 0.92(0.64-1.33) |
| | HWE:1 p-valor | | | 0.21 | 0.016 | | | | | 0.68 | 0.65 |
| rs1022580 | TT | 1438 | 1737 | 96.6 (0.2) | 88.67 (0.34) | 1023 (92.7) | 816 (92.4) | 424 (94.2) | 949 (94.6) | 1 | 1 |
| | GT-GG | 107 | 119 | 95.2 (0.9) | 85.38 (1.27) | 81 (7.3) | 67 (7.6) | 26 (5.8) | 5 (4.5) | 0.77(0.48-1.22) | 0.67(0.45-1.00) |
| | HWE:0.6 p-valor | | | 0.15 | 0.011 | | | | | 0.26 | 0.05 |

Tabla 20. Asociaciones con cintura y OB Central en hombres (H) y mujeres (M) de la población total estudiada en la fase preliminar. N: número de muestras.

5.1.2.2 Asociaciones de los polimorfismos de la CMR y la DMT2

5.1.2.2.1 Resultados de asociaciones de SNPs de CMR y DMT2 en las poblaciones individuales

Hemos analizado los datos de asociación de los estudios de Pizarra, Segovia y Horteiga, y se seleccionaron todos los SNPs que presentaban alguna asociación significativa ($p < 0.05$) en alguna de estas poblaciones, y que tenían la misma tendencia en las restantes, con respecto a su comportamiento con el riesgo de DMT2 y valores de glucemia. Los parámetros estadísticos obtenidos al realizar las asociaciones con DMT2 en las poblaciones independientes se muestran en la tabla 21.

En el Estudio Horteiga, los polimorfismos rs2410718, rs1136224 y rs11205591 se asocian significativamente con el riesgo de padecer DMT2. Los alelos minoritarios de los dos primeros SNPs mostraron un aumento del riesgo de padecer DMT2 y el alelo G del SNP rs11205591 presentó efecto protector contra el riesgo de DMT2. El rs2410718 se asoció de manera similar en la población de Pizarra, lo que refuerza el hallazgo. En la población de Segovia, hemos encontrado una asociación significativa

con el SNP rs4656993, en la que para el alelo A, se ha obtenido menos riesgo de DMT2.

Con respecto al comportamiento de los SNPs asociados con glucemia, el rs17431357 ha mostrado valores más altos de glucemia para el genotipo TT en las tres poblaciones independientes, aunque sólo es significativa la diferencia de glucemias entre genotipos en el estudio Hortega. Estos resultados se muestran en la figura 32. A pesar de que en el estudio Hortega, los SNPs rs4806187 y rs4656993 no han mostrado la misma tendencia que en el resto de poblaciones, esos polimorfismos se han incluido en el estudio posterior en la población conjunta, debido a que las diferencias no son del todo concluyentes. Las asociaciones de las variantes con la glucemia se detallan en la tabla 22.

| Gen SNP | Geno-tipo | ESTUDIO HORTEGA | | | ESTUDIO PIZARRA | | | ESTUDIO SEGOVIA | | |
|---------------------------------|-----------|-----------------|-----------|-----------------------|-----------------|-----------|-----------------------|-----------------|-----------|-----------------------|
| | | No DMT2 (n(%)) | DMT2 (n%) | DMT2 OR(95% CI) | No DMT2 T2(n%) | DMT2 (n%) | DMT2 OR(95% CI) | No DM (n%) | DMT2 (n%) | DMT2 OR(95% CI) |
| COX7C rs2410718 HWE:0.17 | GG | 1225(92) | 92 (84) | 1 | 577(92) | 132(86) | 1 | 859 (92) | 93(93) | 1 |
| | AG | 90 (6.8) | 15 (13) | 2.16/1.15-4.06 | 48 (7.7) | 19(12.5) | 1.88/1.03-3.44 | 67 (7.2) | 7(7) | 0.94/0.41-2.15 |
| | AA | 3 (0.2) | 2 (1.8) | 10.47/1.28-85 | 1 (0.2) | 1 (0.7) | 3.58/0.20-65.5 | 3 (0.3) | 0(0) | 0.00/0.00-NA |
| | p-valor | | | 0.0032 | | | 0.044 | | | 0.71 |
| NDUFS2 rs1136224 HWE:0.26 | AA | 976(74) | 69(64) | 1 | 464(75) | 108(71) | 1 | 721(78.) | 76(73.8) | 1 |
| | AG | 301(23) | 33(30) | 1.73/1.09-2.75 | 141(22) | 36(23.8) | 1.37/0.85-2.21 | 182(20) | 24(23.3) | 1.24/0.75-2.03 |
| | GG | 30(2.3) | 5(4.7) | 2.59/0.89-7.57 | 12(1.9) | 7(4.6) | 2.56/0.81-8.09 | 16(1.7) | 3(2.9) | 1.55/0.43-5.56 |
| | p-valor | | | 0.0077 | | | 0.057 | | | 0.31 |
| NDUFS2 rs4656993 HWE:0.06 | GG | 493(37) | 42(38) | 1 | 237(38) | 62(40.8) | 1 | 330(39) | 55(54.5) | 1 |
| | AG-AA | 824(62) | 67(61) | 0.73/0.48-1.13 | 384(61) | 90(59.2) | 0.79/0.53-1.19 | 503(60) | 6(45.) | 0.54/0.36-0.83 |
| | p-valor | | | 0.16 | | | 0.26 | | | 0.0049 |
| NDUFS5 rs11205591 HWE:0.7 | CC-CG | 1203(91) | 106(98) | 1 | 592(95) | 149(98) | 1 | 813(91) | 92(93.9) | 1 |
| | GG | 113(8.6) | 2(1.8) | 0.19/0.05-0.81 | 30(4.8) | 3(2) | 0.46/0.12-1.67 | 79(8.9) | 6(6.1) | 0.68/0.28-1.66 |
| | p-valor | | | 0.0039 | | | 0.2 | | | 0.38 |

Tabla 21. Asociaciones con DMT2 al analizar los estudios de la fase 1 de forma independiente.

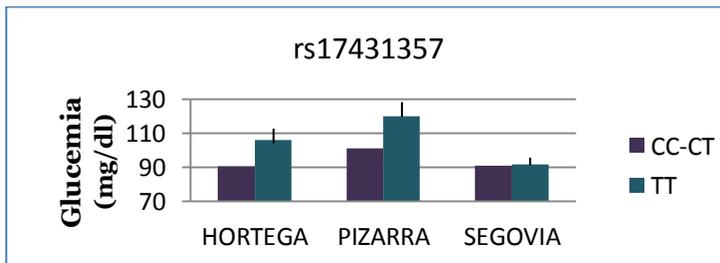


Figura 32. Comportamiento de los genotipos del polimorfismo rs17431357 respecto a las medias de glucemia que presentan, en los tres estudios incluidos en el análisis de asociación (SD es cercana a 0 en los genotipos CC-CT y por eso no es visible en el gráfico)

| Gen SNP | Genotipo | HORTEGA | | PIZARRA | | SEGOVIA | |
|---------------|----------|---------|------------------------------|---------|------------------------------|---------|------------------------------|
| | | N | Glucemia (Media mg/dl±SD) | N | Glucemia (Media mg/dl±SD) | N | Glucemia (Media mg/dl±SD) |
| UQCRQ | GG | 993 | 90.66±0.54 | 491 | 100.25±0.74 | 670 | 89.48±0.57 |
| rs17624157 | AG-AA | 370 | 90.95±0.82 | 204 | 103.58±1.55 | 281 | 91.83±1.24 |
| | p-valor | | 0,74 | | 0.034 | | 0,061 |
| COX6A1 | CC-CT | 1362 | 90.76±0.45 | 689 | 101.13±0.69 | 983 | 90.15±0.53 |
| rs17431357 | TT | 5 | 106.2±12.44 | 3 | 120±14.29 | 6 | 91.67±5.35 |
| | p-valor | | 0.023 | | 0.06 | | 0.67 |
| NDUFS2 | GG | 509 | 89.94±0.61 | 267 | 102.1±1.25 | 375 | 92.01±1.01 |
| rs4656993 | AG-AA | 855 | 91.28±0.62 | 423 | 100.65±0.83 | 541 | 88.96±0.61 |
| | p-valor | | 0.5 | | 0.18 | | 0,009 |
| COX6B1 | AA | 343 | 90.24±1.0 | 197 | 103.28±1.52 | 245 | 91.07 (1.13) |
| rs4806187 | AG | 701 | 91.4±0.62 | 351 | 100.8±0.93 | 472 | 90.14 (0.82) |
| | GG | 322 | 90.01±0.83 | 149 | 99.34±1.3 | 263 | 88.77 (0.83) |
| | p-valor | | 0.99 | | 0.028 | | 0.13 |

Tabla 22. Asociaciones con valores de glucemia en el análisis individual de los estudios de la fase 1.

5.1.2.2.2 Resultados de asociaciones de SNPs de CMR y DMT2 en la población total

En los estudios de asociación de los SNPs de la CMR con la DMT2, en la población conjunta de Pizarra, Hortega y Segovia, hemos encontrado asociaciones estadísticamente significativas con todos los genes que se seleccionaron en el análisis de los estudios individuales, por su asociación con DMT2: COX7C, NDUFS5 y NDUFS2. Los alelos minoritarios del SNP rs2410718 perteneciente a COX7C, y del rs1136224 de NDUFS2, se asociaron con un aumento del riesgo de DMT2, mientras que el alelo G del polimorfismo rs11205591, y el alelo A del rs4656993 (genes NDUFS5 y NDUFS2 respectivamente), con un efecto protector (tabla 23). En todos los casos, hemos obtenido una asociación más significativa en el análisis conjunto, que en el estudio de las poblaciones de forma separada, reforzando dicha asociación. En la figura 33 se muestra el comportamiento de los OR obtenidos al analizar la DMT2 en las distintas muestras.

| Gen SNP | Genotipo | No DMT2 (n(%)) | DMT2 (n(%)) | Riesgo de DMT2 OR(95% CI) p-valor |
|-------------------|----------|----------------|-------------|--------------------------------------|
| COX7C | GG | 2661 (92.6%) | 317 (87.8%) | 1 |
| rs2410718 | AG | 205 (7.1%) | 41 (11.4%) | 1.70 (1.18-2.47) |
| HWE:0.21 | AA | 7 (0.2%) | 3 (0.8%) | 4.34 (1.06-17.80) |
| MAF: 0.199 | p-valor | | | 0,0013 |
| NDUFS5 | CC-CG | 2608 (92.2%) | 347 (96.9%) | 1 |
| rs11205591 | GG | 222 (7.8%) | 11 (3.1%) | 0.37 (0.19-0.72) |
| HWE:0.76 | p-valor | | | 0,0008 |
| NDUFS2 | GG | 1060 (38.2%) | 159 (43.9%) | 1 |
| rs4656993 | AG-AA | 1711 (61.8%) | 203 (56.1%) | 0.74 (0.58-0.93) |
| HWE:0.32 | p-valor | | | 0,0098 |
| MAF: 0.373 | | | | |
| NDUFS2 | AA | 2161 (76%) | 253 (70.1%) | 1 |
| rs1136224 | AG | 624 (21.9%) | 93 (25.8%) | 1.31 (1.01-1.71) |
| HWE:0.12 | GG | 58 (2%) | 15 (4.2%) | 2.20 (1.19-4.06) |
| MAF: 0.188 | p-valor | | | 0,0038 |

Tabla 23. Asociaciones obtenidas con la DMT2 en la población conjunta.

En el caso del estudio de la glucemia, de los cuatro SNPs seleccionados a través del análisis por poblaciones independientes, hemos encontrado una asociación significativa del gen NDUFS2 (rs4656993). Por otro lado, se ha mejorado la significancia de las asociaciones obtenidas para rs17431357 y rs4806187, en el estudio de la población total. No se han obtenido cambios para la asociación del rs17624157 al aumentar el tamaño muestral. En la figura 34 se muestra representada la diferencia entre las medias de la glucemia, de las poblaciones individuales y de la conjunta, para los genotipos con el alelo asociado a valores altos de este SNP. Los resultados detallados de las asociaciones con glucemia se muestran en la tabla 24.

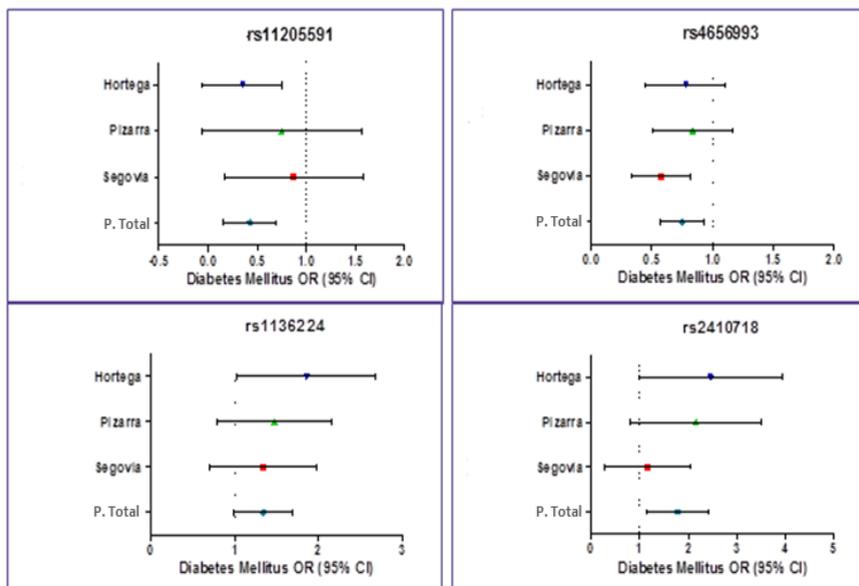


Figura 33. Valores de OR obtenidos para las asociaciones con DMT2 en el análisis de las poblaciones independientes y en de la muestra total.

| Gen SNP | Genotipos | N | Glucemia (Media en mg/dl(SD)) | Diferencias (95% IC) |
|----------------------------------|-----------|------|-------------------------------|------------------------------|
| COX6A1 rs17431357 HWE:0.24 | CC-CT | 3034 | 92.92 (0.32) | 0.00 |
| | TT | 14 | 102.93 (6.14) | 11.21 (2.26 - 20.17) |
| | p-valor | | | 0.014 |
| UQCRQ rs17624157 HWE:0.47 | GG | 2154 | 92.48 (0.36) | 0.00 |
| | AG-AA | 855 | 94.25 (0.68) | 1.47 (0.11 - 2.83) |
| | p-valor | | | 0.034 |
| COX6B1 rs4806187 HWE:0.94 | AA | 785 | 93.77 (0.71) | 0.00 |
| | AG | 1524 | 93.17 (0.45) | -0.47 (-1.95 - 1.00) |
| | GG | 734 | 91.46 (0.56) | -2.08 (-3.80 - -0.36) |
| p-valor | | | 0.019 | |
| NDUFS2 rs4656993 HWE:0.12 | GG | 1151 | 93.44 (0.53) | 0.00 |
| | AG-AA | 1819 | 92.77 (0.41) | -0.89 (-2.15 - 0.37) |
| | p-valor | | | 0.17 |

Tabla 24. Asociaciones obtenidas con la glucemia en la población conjunta.

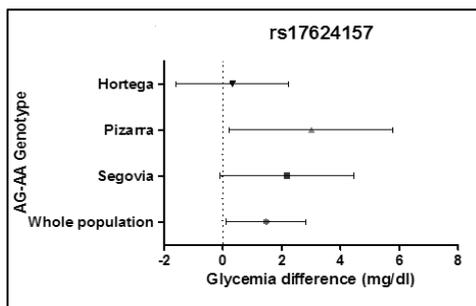


Figura 34. Diferencias de los valores medios de glucemia obtenidos para los genotipos asociados con valores elevados de este parámetro, en las poblaciones estudiadas. Whole population: población total analizada.

5.1.2.2.3 Resultados de asociaciones de SNPs de CMR y DMT2 en subpoblaciones

Hemos agrupado a la población, por un lado, según el género, y por otro, según si presentaban obesidad o no, para su posterior análisis en búsqueda de asociaciones de los SNP, con la DMT2 en esas muestras. De esta forma, podríamos identificar un posible efecto específico en estos grupos.

5.1.2.2.3.1 Resultados de las asociaciones según género

Se han realizado las asociaciones de los polimorfismos de la CMR y la DMT2 en los hombres y las mujeres de la población total de forma independiente. Los resultados detallados se muestran en la tabla 25.

Con respecto a los polimorfismos asociados previamente en la población total, se ha visto que los rs11205591 y rs2410718 se asociaban en hombres con el riesgo de DMT2. Sin embargo, estas asociaciones son más fuertes en la población total (posiblemente por el hecho de que realmente, la asociación depende de toda la muestra y al dividirla se pierde poder estadístico). Otro polimorfismo asociado en hombres con DMT2, es el rs4656993, mejorando esta asociación respecto a la encontrada en la población total, en donde el alelo A se asociaba con menor riesgo de DMT2.

| Gen SNP | Genotipo | NO DMT2 (n/) | | DMT2 (n/) | | OR(95 CI) | |
|----------------------|----------|-----------------|-------------|--------------|---------------|--------------------------|-------------------------|
| | | Hombres | Mujeres | Hombres | Mujeres | Hombres | Mujeres |
| NDUFS5 rs11205591 | CC-CG | 1163 (92.5) | 1445 (91.9) | 184 (97.9) | 163 (95.9) | 1.00 | 1.00 |
| | GG | 94 (7.5) | 127 (8.1) | 4 (2.1) | 7 (4.1) | 0.27 (0.10-0.77) | 0.53 (0.24-1.19) |
| | p-valor | | | | | 0.0031 | 0.098 |
| COX6A1 rs17431357 | CC-CT | 1263 (99.9) | 1573 (99.4) | 185 (98.4) | 170 (99.4) | 1.00 | 1.00 |
| | TT | 1 (0.1) | 10 (0.6) | 3 (1.6) | 1 (0.6) | 21.24(2.11-213.7) | 1.26 (0.14-11.41) |
| | p-valor | | | | | 0.0047 | 0.84 |
| UQCRQ rs17624157 | GG | 902 (72.6) | 1122 (71.7) | 117 (61.6) | 123 (72.3) | 1.00 | 1.00 |
| | AG-AA | 341 (27.4) | 443 (28.3) | 73 (38.4) | 47 (27.6) | 1.85 (1.33-2.58) | 0.83 (0.57-1.21) |
| | p-valor | | | | | 0.0003 | 0.33 |
| UQCRQ rs39830 | TT | 443 (35.3) | 570 (36.4) | 50 (26.7) | 65 (37.4) | 1.00 | 1.00 |
| | CT-CC | 812 (64.7) | 997 (63.6) | 137 (73.3) | 109 (62.6) | 1.64 (1.15-2.35) | 0.87 (0.62-1.23) |
| | p-valor | | | | | 0.0051 | 0.44 |
| COX7A2 rs9360898 | TT | 770 (61.6) | 975 (62.1) | 123 (67.2) | 88 (51.5) | 1.00 | 1.00 |
| | GT-GG | 479 (38.4) | 594 (37.9) | 60 (32.8) | 83 (48.5) | 0.76 (0.54-1.07) | 1.57 (1.12-2.21) |
| | p-valor | | | | | 0.11 | 0.0091 |
| NDUFS2 rs1136224 | AA | 963 (76.5) | 1197 (75.6) | 134 (70.5) | 119 (69.6) | 1.00 | 1.00 |
| | AG | 268 (21.3) | 356 (22.5) | 47 (24.7) | 46 (26.9) | 1.29 (0.89-1.86) | 1.43/0.97-2.11 |
| | GG | 28 (2.2) | 30 (1.9) | 9 (4.7) | 6 (3.5) | 2.02 (0.90-4.57) | 2.59 (1.00-6.74) |
| p-valor | | | | | 0.051 | 0.017 | |
| NDUFS2 rs4656993 | GG | 477 (38.8) | 582 (37.8) | 88 (46.6) | 71 (41) | 1.00 | 1.00 |
| | AG-AA | 751 (61.2) | 959 (62.2) | 101 (53.4) | 102 (59) | 0.63 (0.46-0.87) | 0.77 (0.54-1.08) |
| | p-valor | | | | | 0.0049 | 0.13 |
| COX7C rs2410718 | GG | 1175 (92.2) | 1484 (93) | 164 (87.2) | 153 (88.4) | 1.00 | 1.00 |
| | AG | 98 (7.7) | 107 (6.7) | 21 (11.2) | 20 (11.6) | 1.62 (0.96-2.73) | 1.69 (0.98-2.92) |
| | AA | 2 (0.2) | 5 (0.3) | 3 (1.6) | 0 (0) | 9.68 (1.52-61.80) | 0.00 (0.00-NA) |
| p-valor | | | | | 0.0087 | 0.14 | |

Tabla 25. Asociaciones obtenidas para la población masculina y femenina con la DMT2.

En mujeres, la asociación del SNP rs1136224 con DMT2, es bastante menos significativa que en la población total. El alelo G del SNP rs9360898 se ha asociado con DMT2 solo en la población femenina, mostrando una tendencia totalmente opuesta en hombres.

En la tabla 26 se muestran en detalle las asociaciones obtenidas con glucemia en las poblaciones masculina y femenina. Para los SNPs rs4806187 y rs17431357, las asociaciones son sólo en hombres para el primero, y solamente en mujeres, para el segundo, pero en ambos casos con menor significación que en la muestra total. En hombres, se ha encontrado una nueva asociación del rs7255180 con glucemia.

| Gen SNP | Genotipo | N | | Glucemia (Media en mg/dl(SD)) | | Diferencias (95%IC) | |
|------------------------------------------|----------|---------|---------|----------------------------------|-------------|--------------------------|-------------------------|
| | | Hombres | Mujeres | Hombres | Mujeres | Hombres | Mujeres |
| | | | | | | | |
| COX7A1 rs725180 HWE:1 | CC | 1216 | 1472 | 95.02(0.54) | 91.42(0.43) | 0.00 | 0.00 |
| | CT | 150 | 193 | 92.79(1.31) | 92.62(1.33) | -2.30(-5.39-0.78) | 1.21(-1.18-3.60) |
| | TT | 11 | 6 | 84.55(3.59) | 93.33(5.33) | -10.8(-21.58-0) | 1.49(-11.27-14.24) |
| | p-valor | | | | | 0.028 | 0.31 |
| COX6B1 rs4806187 HWE:0.94 | AA | 364 | 420 | 95.57(1.08) | 92.22(0.93) | 0.00 | 0.00 |
| | AG | 679 | 845 | 95.43(0.73) | 91.36(0.56) | -0.01(-2.34-2.32) | -0.97(-2.84-0.90) |
| | GG | 329 | 404 | 92.19(0.83) | 90.87(0.76) | -3.12(-5.85--0.4) | -1.28(-3.46-0.90) |
| | p-valor | | | | | 0.029 | 0.25 |
| UQCRB rs10504961 HWE:0.43 | CC | 395 | 477 | 96.2(1.12) | 90.47(0.66) | 0.00 | 0.00 |
| | CT-TT | 977 | 1163 | 94.01(0.54) | 92.02(0.52) | -2.43(-4.56--0.3) | 1.79(0.09-3.49) |
| | p-valor | | | | | 0.026 | 0.039 |
| | | | | | | | |
| COX6A1 rs17431357 HWE:0.067 | CC-CT | 1374 | 1658 | 94.68(0.5) | 91.46(0.41) | 0.00 | 0.00 |
| | TT | 3 | 11 | 112(13) | 100.45(7.1) | 15.75(-4.88-36) | 9.99(0.55-19.43) |
| | p-valor | | | | | 0.13 | 0.038 |
| | | | | | | | |

Tabla 26. Asociaciones obtenidas para la población masculina y femenina con los valores de glucemia.

Además, se ha hallado un polimorfismo que se comporta totalmente opuesto en hombres y mujeres. El alelo T del rs10504961, en el primer grupo es protector, y en el otro, da mayor riesgo de padecer la enfermedad. Esto explicaría por qué no se ha hallado asociación en la población total (tabla 26).

5.1.2.2.3.2 Resultados de las asociaciones en población obesa

Con el fin de encontrar los SNPs de la CMR que se asociaban con DMT2 específicamente en la población obesa, los análisis estadísticos de asociación se realizaron seleccionando a los pacientes obesos de la población total. En la tabla 27 se muestran las asociaciones encontradas en esta subpoblación.

Hemos hallado la misma asociación que en la población total del rs4656993, pero con mayor significancia de la asociación en presencia de obesidad. Por otro lado, el genotipo CC del rs10742816 presentaba mayor riesgo de DMT2, si bien el número de portadores de este genotipo es muy bajo.

Al analizar los resultados del estudio de la glucemia, la asociación con el rs17624157 ha sido mucho más significativa en la población obesa que en la total. Los valores para la asociación en la población obesa se muestran en la tabla 28.

| Gen/SNP | Genotipo | DMT2 (n/%) | No DMT2 (n/%) | OR (95 % CI) |
|------------|----------|------------|---------------|----------------------------|
| NDUFS2 | GG | 82 (44.8) | 238 (35.3) | 1 |
| rs4656993 | AG-AA | 101 (55.2) | 436 (64.7) | 0.58 (0.41-0.82) |
| HWE:0.16 | p-valor | | | 0,0021 |
| NDUFS3 | TT-CT | 181 (98.4) | 694 (99.9) | 1 |
| rs10742816 | CC | 3 (1.6) | 1 (0.1) | 19.90 (2.02-196.01) |
| HWE: 0.05 | p-valor | | | 0,0052 |

Tabla 27. Asociaciones halladas en la población obesa para la DMT2.

| Gen SNP | Genotipos | N | Glucemia (Media en mg/dl(SD)) | Diferencias (95% IC) |
|------------|-----------|-----|-------------------------------|-------------------------|
| UQCRQ | GG | 528 | 97.12 (0.85) | 0.00 |
| rs17624157 | AG-AA | 243 | 102.81 (1.62) | 5.69(2.44- 8.94) |
| HWE:0.47 | p-valor | | | 0.0006 |

Tabla 28. Asociación encontrada en la población obesa para la glucemia.

5.1.2.2.3.3 Resultados de las asociaciones en población no obesa

El estudio de las asociaciones de la CMR con la DMT2 en la población no obesa se ha realizado con el fin de evaluar el efecto que produce la obesidad sobre ellas, ya que los resultados obtenidos aquí no están influenciados por su presencia.

Las asociaciones detalladas obtenidas con DMT2, en la población no obesa, se encuentran en la tabla 29. Tres de los cuatro polimorfismos asociados en la población total, también lo hacen con DMT2 en población no obesa. Sin embargo, ninguno de ellos ha mejorado la significancia con respecto a la hallada en las asociaciones de la población conjunta, por lo que el efecto de esos SNPs sobre el riesgo de DMT2 no sería más importante en la población no obesa que en la general. El SNP restante, el rs4656993, no se ha asociado en la subpoblación analizada en este apartado, pero sí en la población obesa, lo que apoya la idea de que la obesidad influye en la acción de este polimorfismo sobre la DMT2.

Con respecto al estudio de la Glucemia en individuos no obesos, se ha encontrado una asociación específica de este grupo, del alelo C del rs1053517 y mayores valores de glucosa en sangre. Al comparar los hallazgos con la población general, se ha visto que sólo se asocia el rs4806187 en no obesos, manteniéndose la misma

significación en ambas poblaciones. Los resultados detallados se muestran en la tabla 30.

| Gen/SNP | Genotipo | No DMT2 (n/%) | DMT2 (n/%) | OR (95 % CI) |
|----------------------------------|----------|---------------|------------|-------------------------|
| NDUFS5 rs11205591 HWE:0.45 | CC-CG | 1923 (92) | 171 (97.2) | 1 |
| | GG | 167 (8) | 5 (2.8) | 0.32 (0.13-0.81) |
| | p-valor | | | 0,0048 |
| NDUFS2 rs1136224 HWE: 0.62 | AA | 1607 (76.8) | 126 (70.8) | 1.00 |
| | AG | 452 (21.6) | 46 (25.8) | 1.43 (0.99-2.06) |
| | GG | 34 (1.6) | 6 (3.4) | 2.62 (1.03-6.68) |
| | p-valor | | | 0,013 |
| NDUFS4 rs535277 HWE: 0.58 | AA | 1460 (70.3) | 107 (60.8) | 1 |
| | AC-CC | 616 (29.7) | 69 (39.2) | 1.51 (1.09-2.10) |
| | p-valor | | | 0,014 |
| COX7C rs2410718 HWE: 0.57 | GG | 1956 (92.5) | 153 (86.4) | 1 |
| | AG | 154 (7.3) | 22 (12.4) | 1.68 (1.02-2.76) |
| | AA | 4 (0.2) | 2 (1.1) | 1.74 (1.12-2.71) |
| | p-valor | | | 0,019 |

Tabla 29. Asociaciones halladas en la población no obesa para la DMT2.

| Gen SNP | Genotipos | N | Glucemia (Media en mg/dl(SD)) | Diferencias (95% IC) |
|---------------------------------|-----------|------|-------------------------------------|------------------------------|
| COX6B1 rs4806187 HWE:0.97 | AA | 570 | 91.61 (0.74) | 0.00 |
| | AG | 1105 | 91.46 (0.47) | -0.31 (-1.86 - 1.24) |
| | GG | 526 | 89.54 (0.59) | -2.21 (-4.03 - -0.40) |
| | p-valor | | | 0.018 |
| NDUFS1 rs1053517 HWE:0.19 | TT | 587 | 89.56 (0.59) | 0.00 |
| | CT-CC | 1603 | 91.5 (0.4) | 1.94 (0.50 - 3.39) |
| | p-valor | | | 0.0085 |

Tabla 30. Asociación encontrada en la población no obesa para la glucemia.

5.1.2.3 Asociaciones de los polimorfismos de la CMR y HTA

5.1.2.3.1 Resultados de asociaciones de SNPs de CMR con HTA en las poblaciones individuales

En los estudios Horteiga, Pizarra y Segovia, hemos encontrado tres polimorfismos relacionados con HTA (tabla 31). Dos de ellos han mostrado asociación significativa

en la población de Horteiga, mientras que el rs4600063, se ha incluido para su estudio posterior, debido a que en el estudio Segovia su p-valor es muy cercano al valor de corte usado para indicar asociación significativa. Además, los tres SNPs mostraron una tendencia similar en las poblaciones restantes.

En la tabla 32 se indican las asociaciones de los SNPs con PAD en las poblaciones individuales. En las tres muestras, el alelo C del rs10742816 mostraba valores de PAD más bajos, obteniéndose asociación significativa para este efecto sólo para el estudio Horteiga. Por el contrario, el alelo C del rs1053517 mostraba valores más altos, asociándose significativamente en las poblaciones de Horteiga y Segovia.

| Gen SNP | Geno tipo | ESTUDIO HORTEIGA | | | ESTUDIO PIZARRA | | | ESTUDIO SEGOVIA | | |
|----------------------------------------------|------------------------|------------------------|------------------------|------------------------------------------------|---------------------|---------------------|---------------------------------|---------------------|---------------------|----------------------------------|
| | | No HTA (n/%) | HFA (n/%) | HTA OR/95% CI | No HTA (n/%) | HTA (n/%) | HFA OR/95% CI | No HTA (n/%) | HFA (n/%) | HTA OR/95% CI |
| UQCRQ rs803224 HWE:0.37 | TT CT-CC p-valor | 671(83.6) 132(16.4) | 500(82.2) 108(17.8) | 1.00 1.46(1.02-2.09) 0.036 | 378(78) 106(21) | 313(80) 74(19.1) | 1.00 1.00(0.64-1.54) 0.99 | 468(80) 114(19) | 360(77) 104(22) | 1.00 1.34(0.97-1.85) 0.079 |
| SDHC rs4600063 HWE:1 | AA AG-GG p-valor | 708(87) 105(12) | 542(87.4) 78(12.6) | 1.00 1.39(0.93-2.06) 0.11 | 429(89) 51(10.6) | 338(88) 44(11.5) | 1.00 1.37(0.77-2.43) 0.28 | 523(88) 67(11.4) | 409(86) 64(13.5) | 1.00 1.42(0.96-2.11) 0.083 |
| UQCRB rs10504961 HWE:0.4 | CC-CT TT p-valor | 647(80) 156(19) | 468(76.7) 142(23.3) | 1.00 1.52(1.10-2.11) 0.012 | 352(72) 131(27) | 276(72) 105(27) | 1.01(0.67-1.51) 1.00 0.96 | 433(77) 126(22) | 366(79) 94(20.4) | 1.00 0.96(0.69-1.32) 0.78 |

Tabla 31. Resultados de las asociaciones con HTA en los estudios Horteiga, Pizarra y Segovia.

Los detalles de los hallazgos para PAS en las poblaciones individuales se indican en la tabla 33.

| Gen SNP | Genotipo | ESTUDIO HORTEIGA | | | ESTUDIO PIZARRA | | | ESTUDIO SEGOVIA | | |
|---------------------------------------------------------------|------------------------|------------------|----------------------------|-------------------------------------------------------|-----------------|----------------------------|----------------------------------|-----------------|----------------------------|------------------------------------------------|
| | | N | PAD (mmHg/SD) | Diferencias (95% IC) | N | PAD (mmHg/SD) | Diferencias (95% IC) | N | PAD (mmHg/SD) | Diferencias (95% IC) |
| NDUFS3 rs10742816 HWE:0.85 MAF: | TT CT-CC p-valor | 959 185 | 78.74(0.34) 77.28(0.71) | 0.00 -1.58(-3.07-- 0.1) 0.038 | 556 125 | 76.22(0.5) 75.56(1.05) | 0.00 -1.18(-3.20-0.8) 0.25 | 785 147 | 77.86(0.37) 76.45(0.8) | 0.00 -0.89(-2.58-0.8) 0.3 |
| NDUFS1 rs1053517 HWE:1 MAF:0.91 | TT CT-CC p-valor | 321 807 | 77.51(0.52) 78.95(0.37) | 0.00 1.46(0.23-2.68) 0.02 | 154 521 | 75.92(1.01) 76.14(0.51) | 0.40(-1.47-2.2) 0.00 0.68 | 253 642 | 76.89(0.59) 77.68(0.42) | 0.00 1.47(0.08-2.86) 0.039 |

Tabla 32. Asociaciones obtenidas con PAD en los estudios Horteiga, Pizarra y Segovia.

| SNP | Geno tipo | ESTUDIO HORTEGA | | | ESTUDIO PIZARRA | | | ESTUDIO SEGOVIA | | |
|-----------------------------|------------------------|-----------------|------------------------------|-----------------------------------|-----------------|------------------------------|-------------------------------------------------|-----------------|------------------------------|--------------------------------------------------|
| | | N | PAS (mmHg/SD) | Diferencias (95% IC) | N | PAS (mmHg/SD) | Diferencias (95% IC) | N | PAS (mmHg/SD) | Diferencias (95% IC) |
| rs39830 HWE:0.5 | TT-CT CC p-valor | 955 177 | 127.17(0.65) 127.14(1.41) | 0.00 0.60(-1.94-3.14) 0.64 | 567 115 | 125.15(0.82) 120.98(1.61) | 0.00 -3.57(-6.71-0.4) 0.025 | 730 159 | 122.54(0.64) 121.48(1.25) | 0.00 -0.34(-2.92-2.24) 0.8 |
| rs1133322 HWE:0.4 | TT CT-CC p-valor | 309 819 | 126.98(1.15) 127.01(0.68) | 0.00 -0.47(-2.53-1.60) 0.66 | 187 492 | 124.02(1.45) 124.65(0.86) | 0.00 0.81(-1.83-3.45) 0.55 | 250 615 | 119.97(1.2) 122.98(0.67) | 0.00 2.35(0.03-4.68) 0.048 |
| rs7255180 HWE:0.7 | CC-CT TT p-valor | 112 97 | 127.04(0.59) 125(4.27) | 0.00 -1.07(-12.8-10.6) 0.86 | 677 2 | 124.38(0.74) 123(23) | 0.00 -6.32(-28.0-15.4) 0.57 | 906 5 | 122.73(0.58) 108.6(6.43) | 0.00 -17.6(-31.7--3.6) 0.014 |

Tabla 33. Asociaciones obtenidas con PAS en los estudios Horteiga, Pizarra y Segovia.

5.1.2.3.2 Resultados de asociaciones de SNPs de CMR con HTA en la población total

Se ha realizado el análisis de asociación con hipertensión arterial, de los SNPs seleccionados a partir de los resultados de las poblaciones individuales, en la población compuesta por la suma de ellas.

Se ha obtenido asociación significativa para dos de los tres SNPs seleccionados, el rs4600063 y el rs803224. En ambos casos, sus alelos minoritarios se han relacionado con mayor riesgo de HTA. Por el contrario, el rs10504961 no ha llegado a tener una asociación con HTA al aumentar el tamaño muestral. Los resultados se describen en la tabla 34.

| Gen SNP | Genotipo | No HTA (n(%)) | HTA (n(%)) | Riesgo de HTA OR(95% CI) |
|-----------------------------------------------|------------------------|---------------------------|---------------------------|-----------------------------------------------|
| SDHC rs4600063 HWE:0.66 | AA AG-GG p-valor | 1674 (88.2) 224 (11.8) | 1275 (87.3) 185 (12.7) | 1 1.39 (1.08-1.78) 0.0077 |
| UQCRQ rs803224 HWE:0.21 | TT CT-CC p-valor | 1531 (81.2) 354 (18.8) | 1159 (80.3) 284 (19.7) | 1 1.29 (1.05-1.59) 0.0098 |
| UQCRB rs10504961 HWE:0.40 | CC-CT TT p-valor | 1442 (77.5) 419 (22.5) | 1100 (76.7) 335 (23.3) | 1 1.14 (0.94-1.39) 0.14 |

Tabla 34. Resultados de las asociaciones encontradas con HTA en la población conjunta.

Al analizar como variable a PAD, 2 SNPs seleccionados para su estudio en la población total se han asociado de forma significativa, confirmando así su posible relación con este parámetro. A pesar de no dar significativa su asociación, el polimorfismo rs1133322, se ha incluido para su posterior réplica por la cercanía de su p-valor al punto de corte utilizado. Los detalles de estas asociaciones se muestran en la tabla 35.

En el caso de PAS, de los SNPs estudiados, no hemos obtenido ninguna asociación significativa para ninguno de ellos, en la población conjunta, por lo que no se muestran resultados de este análisis.

| Gen dbSNP | Genotipo | Número de pacientes (N) | PAD (Media en mg/dl(SD)) | Diferencias (95% IC) |
|------------------------------------------------|----------|-------------------------|--------------------------|------------------------------|
| NDUFS3 rs10742816 HWE:0.16 | TT | 2300 | 77.93 (0.23) | 0.00 |
| | CT-CC | 457 | 76.63 (0.48) | -1.38 (-2.38 - -0.38) |
| | p-valor | | | 0.0068 |
| NDUFS1 rs1053517 HWE:1 | TT | 728 | 77.06 (0.38) | 0.00 |
| | CT-CC | 1970 | 77.89 (0.25) | 0.99 (0.14 - 1.83) |
| | p-valor | | | 0.022 |
| COX5A rs1133322 HWE: 0.07 | TT | 745 | 76.95 (0.39) | 0.00 |
| | CT-CC | 1926 | 77.89 (0.24) | 0.73 (-0.11 - 1.56) |
| | p-valor | | | 0.088 |

Tabla 35. Resultados de las asociaciones encontradas con PAD en la población conjunta.

5.1.2.3.3 Resultados de asociaciones de SNPs de CMR y HTA en subpoblaciones

Para detectar posibles interacciones entre la obesidad y el género, y los efectos de los polimorfismos sobre el desarrollo de la DMT2, hemos dividido a la población total en subgrupos, según esos 2 factores, y hemos analizado las asociaciones con los SNPs y la HTA.

5.1.2.3.3.1 Resultados de las asociaciones según género

En el análisis de los polimorfismos en la **población femenina**, se ha encontrado una asociación significativa con el rs1136224 y la HTA (tabla 36), mientras que para los

SNPs relacionados con esa misma enfermedad en la población total, no se ha visto ninguna. Por su parte, se han hallado tres asociaciones con PAD (tabla 37), de las cuales una corresponde al SNP rs10742816, previamente asociado a este parámetro en la población total. Sin embargo, la asociación es más fuerte en la población general, lo que podría ayudar a descartar la influencia del género en esa asociación.

| Gen dbSNP | Genotipo | No HTA (n(%)) | HTA (n(%)) | Riesgo de HTA OR(95% CI) |
|---------------|----------|---------------|------------|--------------------------|
| NDUFS2 | AA-AG | 1049 (98.5%) | 739 (97%) | 1.00 |
| rs1136224 | GG | 16 (1.5%) | 23 (3%) | 2.89 (1.23-6.75) |
| HWE:1 | p-valor | | | 0.013 |

Tabla 36. Resultados de las asociaciones encontradas con HTA en la población femenina.

En el estudio de PAS, se han visto dos SNPs asociados en mujeres, en los cuales sus alelos minoritarios han mostrado valores más altos de PAS. Los detalles de estos polimorfismos se indican en la tabla 38. Otro hallazgo interesante es la asociación de PAS y PAD con el polimorfismo rs17431357, mostrando valores más altos en ambos casos para el alelo T. Este hecho serviría para reforzar la idea de que este SNP afectaría a la presión arterial en la población femenina. Por otro lado, el SNP rs1136224 se ha asociado con HTA y PAS, lo que podría sugerir que su efecto sobre el riesgo de HTA en mujeres se deba a su posible acción sobre la presión sistólica.

| Gen dbSNP | Genotipo | Número de pacientes (N) | PAD (Media en mg/dl(SD)) | Diferencias (95% IC) |
|------------------|----------|-------------------------|--------------------------|------------------------------|
| SDHC | TT | 1258 | 76.2 (0.31) | 0.00 |
| rs13374037 | AT | 178 | 77.4 (0.84) | 1.58 (0.01 - 3.15) |
| HWE:0,017 | AA | 13 | 81.31 (4.16) | 5.54 (0.07 - 11.01) |
| | | | | 0.0082 |
| NDUFS3 | TT | 1225 | 76.62 (0.32) | 0.00 |
| rs10742816 | CT-CC | 237 | 74.73 (0.68) | -1.76 (-3.15 - -0.37) |
| HWE:0,25 | p-valor | | | 0.013 |
| COX6A1 | CC | 1270 | 76.07 (0.3) | 0.00 |
| rs17431357 | CT-TT | 178 | 78.28 (0.9) | 1.70 (0.13 - 3.26) |
| HWE:0,059 | p-valor | | | 0.034 |

Tabla 37. Resultados de las asociaciones encontradas con PAD en la población femenina.

| Gen dbSNP | Genotipo | Número de pacientes (N) | PAS (Media en mg/dl(SD)) | Diferencias (95% IC) |
|------------------------------------------|----------|-------------------------|--------------------------|----------------------------|
| COX6A1 rs17431357 HWE:0,059 | CC | 1270 | 121.25 (0.54) | 0.00 |
| | CT-TT | 178 | 126.29 (1.79) | 3.90 (1.23 - 6.58) |
| | p-valor | | | 0.0043 |
| NDUFS2 rs1136224 HWE:1 | A/A-A/G | 1418 | 121.69 (0.53) | 0.00 |
| | GG | 27 | 130.41 (4.84) | 9.17 (2.69 - 15.64) |
| | p-valor | | | 0.0056 |

Tabla 38. Resultados de las asociaciones encontradas con PAS en la población femenina.

En la **población masculina**, se han asociado las tres variantes que se habían seleccionado para su estudio en HTA en la población total. Sin embargo, sólo se ha obtenido una asociación más significativa en hombres que en la población total, con el SNP rs4600063. Los resultados de las asociaciones se muestran en la tabla 39.

| Gen dbSNP | Genotipo | No HTA (n%) | HTA (n%) | Riesgo de HTA OR(95% CI) p-valor |
|---------------------------------------|----------|-------------|------------|-------------------------------------|
| SDHC rs4600063 HWE:1 | AA | 741 (89.6) | 601 (86.6) | 1 |
| | AG-GG | 86 (10.4) | 93 (13.4) | 1.68 (1.18-2.40) |
| | p-valor | | | 0.0042 |
| UQCRQ rs803224 HWE:0.21 | TT | 667 (81.4) | 546 (79.2) | 1 |
| | CT-CC | 152 (18.6) | 143 (20.8) | 1.40 (1.05-1.87) |
| | p-valor | | | 0.022 |
| UQCRB rs10504961 HWE:0.4 | CC-CT | 637 (78.5) | 526 (76.6) | 1 |
| | TT | 175 (21.6) | 161 (23.4) | 1.35 (1.02-1.78) |
| | p-valor | | | 0.036 |

Tabla 39. Resultados de las asociaciones encontradas con PAS en la población masculina.

Al analizar la variable PAD, se ha encontrado una asociación más significativa con el polimorfismo rs1053517, comparada con la población total. Esto podría deberse a que el efecto de la asociación sobre la presión se ve influenciado fuertemente por el sexo. Los resultados detallados de este polimorfismo se muestran en la tabla 40. Ningún SNP se ha asociado con PAS en esta subpoblación, de la misma forma que ha ocurrido con esta variable en la población total.

| Gen dbSNP | Genotipo | Número de pacientes (N) | PAD (Media en mg/dl(SD)) | Diferencias (95% IC) |
|-----------------|----------|-------------------------|--------------------------|---------------------------|
| NDUFS1 | TT | 339 | 77.73 (0.52) | 0.00 |
| rs1053517 | CT-CC | 924 | 79.85 (0.34) | 2.18 (0.96 - 3.41) |
| HWE:0,47 | p-valor | | | 0.0005 |

Tabla 40. Resultados de las asociaciones encontradas con PAD en la población masculina.

5.1.2.3.2.2 Resultados de las asociaciones en población obesa

Los resultados de las asociaciones con HTA halladas en la población obesa, se muestran en la tabla 41. Al comparar los hallazgos con los obtenidos en la población total, se ha visto que los alelos de los SNPs asociados con mayor riesgo de HTA, mostraban una asociación menos fuerte en obesos.

| Gen dbSNP | Genotipo | No HTA (n(%)) | HTA (n(%)) | Riesgo de HTA OR(95% CI) |
|-----------------|----------|---------------|------------|--------------------------|
| SDHC | AA | 274 (93.2) | 566 (89) | 1.00 |
| rs4600063 | AG-GG | 20 (6.8) | 70 (11) | 1.76 (1.01-3.08) |
| HWE:0,66 | p-valor | | | 0.041 |
| UQCRQ | TT | 242 (84.9) | 509 (80.8) | 1.00 |
| rs803224 | CT-CC | 43 (15.1) | 121 (19.2) | 1.67 (1.09-2.56) |
| HWE:0.21 | p-valor | | | 0.016 |

Tabla 41. Resultados de las asociaciones encontradas con HTA en la población obesa.

No se ha hallado ninguna asociación con PAD de los polimorfismos analizados en la población obesa.

Al analizar los polimorfismos frente a PAS, se ha encontrado una relación significativa entre el alelo T del rs17431357 y valores altos de ese parámetro. Esta asociación no se ha visto en la población total, pero sí en mujeres, donde se vio una asociación más significativa que en obesos. Los resultados se muestran en la tabla 42.

| Gen dbSNP | Genotipo | Número de pacientes (N) | PAS (Media en mg/dl(SD)) | Diferencias (95% IC) |
|----------------------------------|----------|-------------------------|--------------------------|---------------------------|
| COX6A1 rs17431357 HWE:0,55 | CC | 523 | 132.59 (0.8) | 0.00 |
| | CT-TT | 82 | 139.33 (2.3) | 5.05 (1.22 - 8.87) |
| | p-valor | | | 0.0099 |

Tabla 42. Asociación encontrada con PAS en la población obesa.

5.1.2.3.3 Resultados de las asociaciones en población no obesa

En los individuos no obesos, no se han encontrado asociaciones significativas entre los polimorfismos analizados y las variables HTA y PAS. Sin embargo, sí se han asociado con PAD los mismos SNPs que en la población total, el rs10742816 y el rs1053517. Además, se ha encontrado una asociación muy poco significativa con el SNP rs11205591 y esa variable, sólo al estudiar los individuos no obesos. Los resultados de estas asociaciones se muestran en la tabla 43.

| Gen dbSNP | Genotipo | Número de pacientes (N) | PAD (Media en mg/dl(SD)) | Diferencias (95% IC) |
|----------------------------------|----------|-------------------------|--------------------------|------------------------------|
| NDUFS3 rs10742816 HWE:0,25 | TT | 1746 | 76.7 (0.25) | 0.00 |
| | CT-CC | 342 | 75.23 (0.54) | -1.41 (-2.52 - -0.29) |
| | p-valor | | | 0.013 |
| NDUFS1 rs1053517 HWE:1 | TT | 546 | 75.69 (0.42) | 0.00 |
| | CT-CC | 1502 | 76.69 (0.27) | 0.97 (0.02 - 1.91) |
| | p-valor | | | 0.045 |
| NDUFS5 rs11205591 HWE:1 | CC-CG | 1898 | 76.31 (0.24) | 0.00 |
| | GG | 157 | 77.78 (0.79) | 1.62 (0.06 - 3.19) |
| | p-valor | | | 0.042 |

Tabla 43. Resultados de las asociaciones encontradas con PAD en la población no obesa.

5.2 Fase 2: Identificación de variaciones genéticas en los genes asociados

En el paso de detección de nuevas variantes que se asociaran con Diabetes y otros desórdenes relacionados, éstas se han buscado en los genes asociados con DMT2 y OB, que ha sido donde se han encontrado más asociaciones significativas.

Luego de haber seleccionado los genes asociados de manera más fuerte con DMT2 y OB, éstos se han secuenciado en 100 muestras elegidas según sus genotipos y sus fenotipos para tener controles y casos. Hemos comprobado que los polimorfismos hallados cumplían el equilibrio de Hardy Weinberg, y comparado la frecuencia entre casos y controles, y se han realizado los análisis de asociación para ver si se hallaba relación con DMT2 y OB, con los SNPs que mostraron distintas frecuencias alélicas. Se han secuenciado las regiones exónicas e intrónicas contiguas de los genes SDHB, SDHD, NDUFS1 y NDUFS5. Todos estos genes se han asociado al IMC en la población general. El primer y el tercer gen se asociaron significativamente a IMC analizando sólo a las mujeres, además de asociarse en la población general. NDUFS1 se ha asociado también de manera significativa con el riesgo de padecer OB, y NDUFS5 al riesgo de DMT2.

Estos genes se secuenciaron en pacientes obesos y controles sin la enfermedad cuando se relacionaban con IMC y OB, y el gen NDUFS5 se secuenció en pacientes diabéticos y obesos versus controles sin ninguna de las enfermedades, debido a su asociación con la DMT2 y al IMC. Los SNPs de los cuales se han encontrado diferencias en las frecuencias alélicas se muestran en la tabla 44.

Analizando el gen SDHB, en el rs2746462 (p.Ala6=), todos los pacientes obesos presentaban el alelo A en homocigosis, mientras que el alelo mayoritario sólo se encontraba en los individuos controles. En el gen NDUFS5, se han encontrado cuatro SNPs y en tres de ellos el alelo minoritario estaba predominantemente en los casos. El alelo A del rs6981 sólo se observó en los casos. El alelo menos frecuente del SNP rs77655991, perteneciente al gen NDUFS1, sólo estaba presente en los enfermos.

| <u>GEN</u> | MAF ncbi / MAF M | SNP | Ubicación | Observación |
|---------------|---------------------|------------|----------------|--------------------------------------------------------------|
| <u>SDHB</u> | 0.052/0.051 | rs2746462 | E1 (p.Ala6=) | Todos los casos presentan el polimorfismo en homocigosis(AA) |
| <u>NDUFS5</u> | 0.462/0.136 | rs10157617 | 5' del Prom | Alelo T presente mayoritariamente en los casos |
| | 0.461/0.150 | rs1984600 | Prom | Alelo G presente mayoritariamente en los casos |
| | 0.462/0.140 | rs1122472 | E1 | Alelo G presente mayoritariamente en los casos |
| | 0.015/0.035 | rs6981 | 3'UTR | Alelo A presente sólo en los casos |
| <u>NDUFS1</u> | 0.023/0.179 | rs77655991 | intrónico | Alelo A presente sólo en los casos |
| | 0.38/0.357 | rs4147718 | intrónico | Mayor presencia del alelo T en los controles |
| | 0.32/0.238 | rs1801318 | E12(p.Arg164=) | Mayor presencia del alelo A en los controles |
| | 0.380/0.385 | rs4147729 | intrónico | Mayor presencia de homocigotos T en los controles |

Tabla 44. Descripción de los polimorfismos en los cuales se obtuvo diferencias entre sus frecuencias alélicas en el estudio de los casos y controles. M: muestra analizada. Prom: promotor del gen, E1: exón 1 del gen, E12: exón 12 del gen.

5.3 Fase 3: Estudio de las variaciones identificadas

Una vez finalizada la secuenciación masiva en los estudios Pizarra, Segovia y Horteiga, de las 9 variantes identificadas en la fase 2, hemos procedido al análisis bioinformático de los resultados. En primer lugar, se ha comprobado que todos los parámetros que deben evaluarse en la carrera, hayan dado según las especificaciones de Illumina para una carrera de este tipo. Luego de obtener los archivos de secuencias en formato FastQ, se realizó el control de calidad de los datos obtenidos a partir de dichos archivos. Después de realizar el control de calidad de la secuenciación, se han calculado las coberturas de los amplicones secuenciados.

Finalmente, se ha realizado el protocolo (pipeline) bioinformático para obtener los genotipos de los polimorfismos de interés, para posteriormente, proceder a la búsqueda de asociaciones con OB, DMT2 y HTA en la muestra.

Se han tomado como criterios de aceptación de los genotipos obtenidos en la secuenciación, un número mínimo de lecturas para cada SNP y para cada muestra, de 30x, y un Qscore, mayor a 30 (Q30) para cada base. Se dieron por válidos los SNPs con un porcentaje mayor al 80% de las muestras secuenciadas que cumplen con los criterios mencionados. No se han excluido ninguno de los nueve SNPs para los estudios de asociación. Tampoco se ha excluido ningún SNPs al evaluar el equilibrio de Hardy-Weinberg, todos lo han cumplido.

5.3.1 Control de calidad de la secuenciación

-Comprobación de las especificaciones de la carrera en el Miseq

El Miseq Control Software es el encargado de calcular los parámetros necesarios para la evaluación del correcto funcionamiento de la carrera. Un resumen de los valores obtenidos en esta secuenciación, para cada tipo de lectura (read) por separado, se muestra en la figura 35.

Al comparar los valores de nuestra carrera con los definidos por Illumina, la cantidad de clusters que han pasado el filtro de calidad (631 K/mm^2) está por debajo de lo esperado. A pesar de esto, las lecturas que han pasado el filtro de calidad ($12,94 \text{ M} \times 2$) están dentro del rango establecido, y el porcentaje de bases con un Qscore mayor a Q30 es mayor al 80%. Por otro lado, el rendimiento obtenido (4 Gb) no ha llegado al mínimo esperado.

El porcentaje medio de alineamiento esperado para el Phix, secuenciado en paralelo en la carrera, era 5, y se ha obtenido 5.85. El porcentaje medio de la tasa de error, calculado según el alineamiento del PhiX, ha sido correcto porque fue menor a 2.

| Read 1 | | | | | | | | | | | |
|------------|-------|-----------------|----------------|------------------|-----------|--------------|--------------|-----------|------------------|---------------|----------------|
| Lane | Tiles | Density (K/mm2) | Cluster PF (%) | Phas/Prephas (%) | Reads (M) | Reads PF (M) | % \geq Q30 | Yield (G) | Cycles Err Rated | Aligned (%) | Error Rate (%) |
| 1 | 28 | 817 +/- 35 | 77.27 +/- 6.43 | 0.013 / 0.026 | 16.84 | 12.94 | 89.5 | 1.9 | 150 | 5.98 +/- 0.42 | 0.76 +/- 0.23 |
| Read 2 (I) | | | | | | | | | | | |
| Lane | Tiles | Density (K/mm2) | Cluster PF (%) | Phas/Prephas (%) | Reads (M) | Reads PF (M) | % \geq Q30 | Yield (G) | Cycles Err Rated | Aligned (%) | Error Rate (%) |
| 1 | 28 | 817 +/- 35 | 77.27 +/- 6.43 | 0.000 / 0.000 | 16.84 | 12.94 | 78.7 | 0.1 | 0 | 0.00 +/- 0.00 | 0.00 +/- 0.00 |
| Read 3 (I) | | | | | | | | | | | |
| Lane | Tiles | Density (K/mm2) | Cluster PF (%) | Phas/Prephas (%) | Reads (M) | Reads PF (M) | % \geq Q30 | Yield (G) | Cycles Err Rated | Aligned (%) | Error Rate (%) |
| 1 | 28 | 817 +/- 35 | 77.27 +/- 6.43 | 0.000 / 0.000 | 16.84 | 12.94 | 81.0 | 0.1 | 0 | 0.00 +/- 0.00 | 0.00 +/- 0.00 |
| Read 4 | | | | | | | | | | | |
| Lane | Tiles | Density (K/mm2) | Cluster PF (%) | Phas/Prephas (%) | Reads (M) | Reads PF (M) | % \geq Q30 | Yield (G) | Cycles Err Rated | Aligned (%) | Error Rate (%) |
| 1 | 28 | 817 +/- 35 | 77.27 +/- 6.43 | 0.000 / 0.014 | 16.84 | 12.94 | 80.3 | 1.9 | 150 | 5.72 +/- 0.43 | 0.83 +/- 0.20 |

Figura 35. Resumen de los parámetros calculados durante la carrera de secuenciación para comprobar el correcto funcionamiento. Todos los valores se muestran para los cuatro pasos de la secuenciación (lecturas 1, 2, 3 y 4).

-Evaluación de los archivos FastQ

Con el fin de comprobar la calidad de las secuencias obtenidas en la secuenciación realizada, se ha utilizado el programa FastQC para calcular los diferentes parámetros necesarios para realizar su control en todas las muestras. Se han seleccionado los controles específicos para el modo de secuenciación por amplicones. Los valores obtenidos se resumen en la tabla 45.

| Parámetro | Muestras que pasan | Muestras con alerta | Muestras que fallan |
|-----------------------------|--------------------|---------------------|---------------------|
| Per base N content | 3001 | 2 | 7 |
| Per base sequence quality | 484 | 8 | 2518 |
| Per sequence quality scores | 2999 | 4 | 7 |

Tabla 45. Parámetros del control de calidad que son específicamente útiles en la carrera de secuenciación de amplicones realizada en esta fase.

A continuación se muestran como ejemplo, los gráficos de los parámetros obtenidos para una muestra en particular (Muestra 2 del Estudio Hortega) correspondientes a la lectura 1, después del análisis en FastQC:

-Contenido de N por base: se ha comprobado que el porcentaje de bases ambiguas (N) en cada base secuenciada sea igual a cero. En la gráfica de la figura 36 se puede ver que es cero a lo largo de toda la lectura.

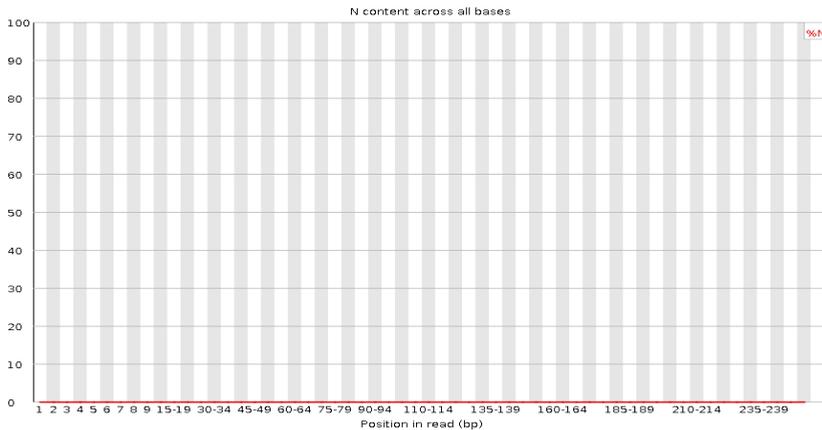


Figura 36. Contenido de N por base para la muestra 2 del Estudio Hortega.

-Calidad de secuencia por base: Se ha observado que el Qscore medio era mayor a 30 hasta aproximadamente la posición 185 de la lectura, y que conforme iba aumentando la longitud de la secuencia, iba bajando. Esta disminución del Qscore fue debida a que el rendimiento de los reactivos baja a medida que pasa el tiempo que dura la secuenciación. Recientemente la casa comercial Illumina ha enviado un comunicado, en el que ha comentado que los reactivos que se han utilizado en esta carrera han estado mostrando disminuciones de calidad al acercarse al final de la secuenciación en mayor medida que lo normal, y que estaban tomando medidas para solucionar el problema. En la figura 37 se grafica la calidad de secuencia por base de una muestra.

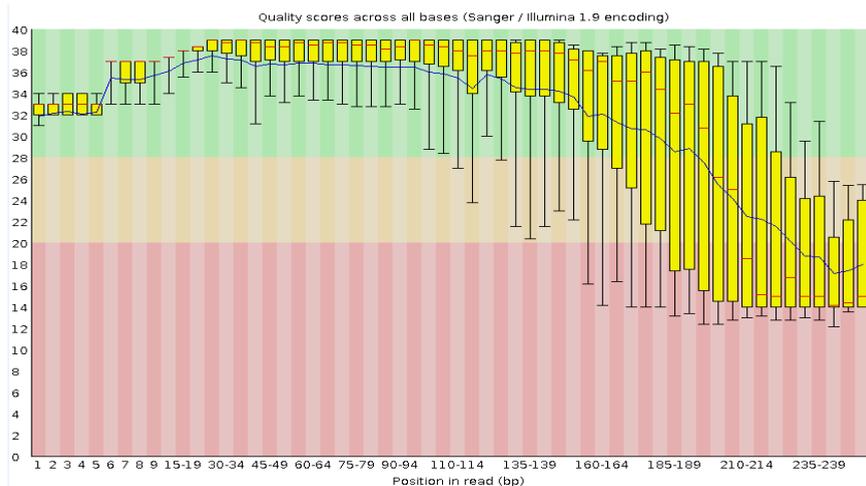


Figura 37. Calidad de la secuencia por base para la muestra 2 del Estudio Hortega. La línea roja representa la mediana y la línea azul la calidad media de secuenciación. -Scores de calidad por secuencia.

-Scores de calidad por secuencia: con este parámetro se ha podido ver que el Qscore tiene una distribución correcta en todas las secuencias, se comportan todas se manera similar (figura 38).

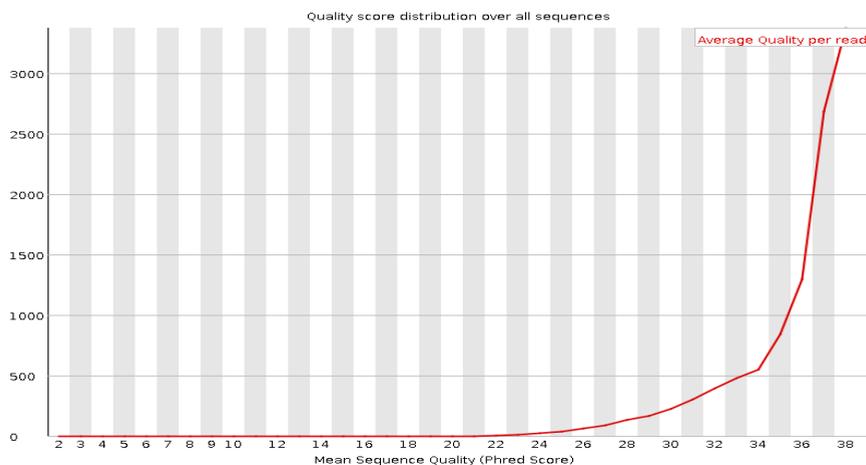


Figura 38. Scores de calidad por secuencia para la muestra 2 del Estudio Hortega.

Por tanto, al revisar los parámetros de control de calidad de la secuenciación, se ha podido concluir que las secuencias obtenidas presentaban las calidades correctas.

5.3.2 Cobertura de los amplicones secuenciados

Después de realizar el control de calidad de la secuenciación, se han analizado los datos de cobertura de las secuencias obtenidas.

La cobertura obtenida para todos los SNPs ha sido mayor respecto al valor de corte de 30x, en la mayoría de las muestras, por lo que se han obtenido unas coberturas adecuadas para el uso de los genotipos obtenidos en la secuenciación, en los análisis de asociación posteriores. Los valores medios de cobertura para cada amplicón se detallan en la figura 39.

5.3.3 Asociaciones de los polimorfismos seleccionados en la fase 2 con DMT2, OB y HTA

Se han realizado los estudios de asociación de los nueve polimorfismos identificados en la secuenciación de controles y casos de la fase 2, en las poblaciones Pizarra, Hortega y Segovia. Se ha calculado el equilibrio de Hardy Weinberg, y se ha visto que todos los polimorfismos lo cumplían. El rs4147729 ha salido monomórfico en esta población, por lo que no se ha podido incluir en el análisis. Todos los SNPs secuenciados han presentado un porcentaje de genotipado mayor del 80%, por lo cual se han incluido en su estudio posterior.

5.3.3.1 Asociaciones con OB de los polimorfismos seleccionados en la población de la fase 1

En la población total analizada, se han hallado asociaciones significativas con el alelo G del rs2746462 y valores bajos de IMC y Cintura, como así también con menor riesgo de OB. En la población femenina, este polimorfismo se ha relacionado con IMC y Cintura, mostrando una asociación más fuerte en el caso del IMC, lo que podría deberse a que el género estuviera interactuando con esta asociación. El SNP rs77655991 se ha asociado sólo en mujeres, con el riesgo de OB, lo que podría

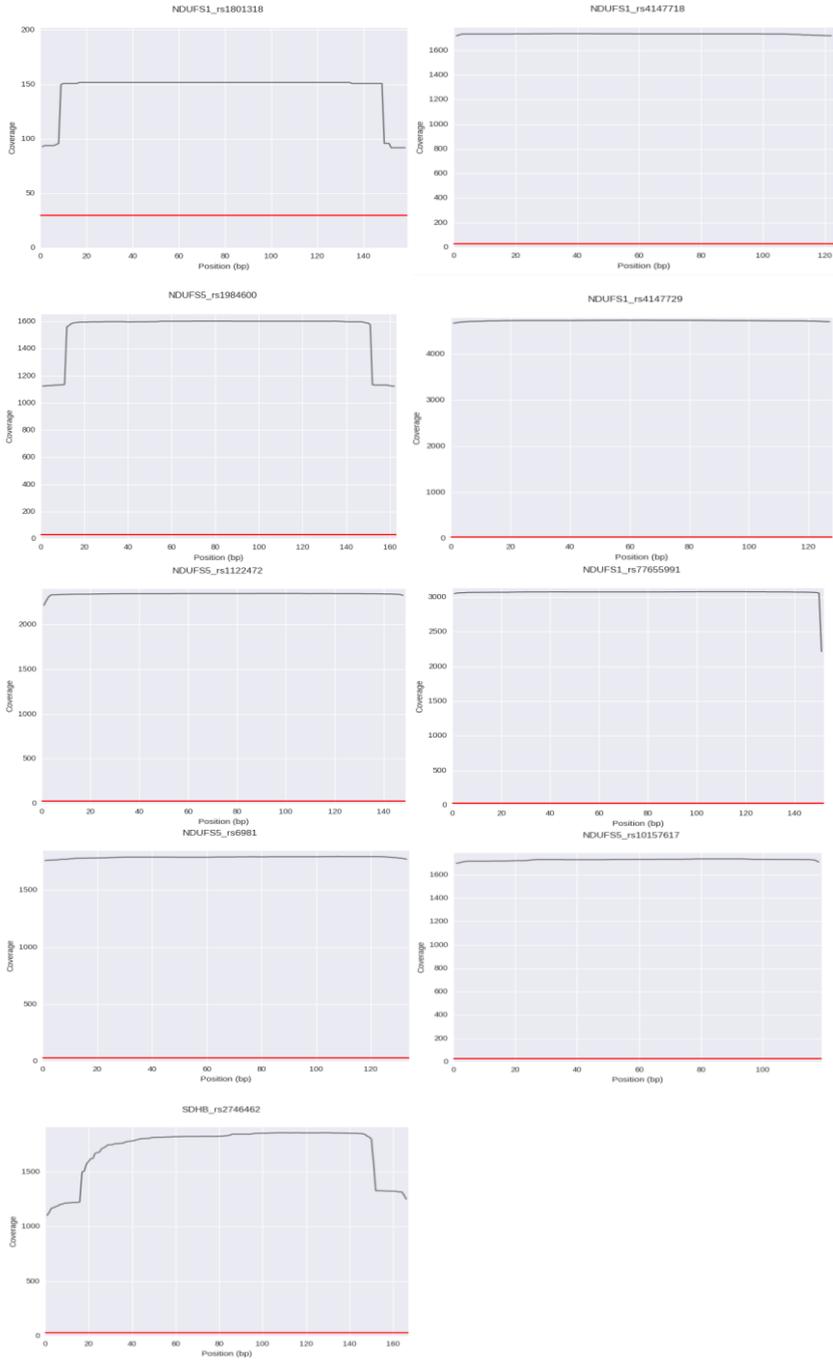


Figura 39. Coberturas de los 9 amplicones secuenciados a lo largo de todo el fragmento.

significar también que el género es un factor que influiría en la aparición de este efecto. Los resultados detallados de las asociaciones se muestran en la tabla 46.

En el caso de la población masculina, no se han encontrado asociaciones con ningún parámetro relacionado con obesidad.

| Gen dbSNP | Genotipos | N | IMC (Kg/m ²) | Cintura (cm) | No OB | | | | | |
|-----------------|-----------|------|--------------------------|--------------|-------------|------------|-----------------|------------|-------------|-----------------|
| | | | | | OB (n%) | OR(95% CI) | No OB C. (n%) | OB C. (n%) | OR(95% CI) | |
| SDHB | TT | 5202 | 27.42±0.08 | 9.7±0.23 | 2387 (93.2) | 815 (94.8) | 1 | 1798 (93) | 1422 (94.6) | 1 |
| rs2746462 | GT-GG | 218 | 26.56±0.28 | 91.1±0.86 | 173(6.8) | 45(5.2) | 0.76(0.54-1.07) | 135(7) | 81 (5.4) | 0.73(0.54-0.99) |
| HWE:0.26 | p-valor | | 0.0062 | 0.0100 | | | 0.11 | | | 0.0430 |
| MUJERES | | | | | | | | | | |
| SDHB | TT | 1746 | 27.38±0.12 | 90.84±0.35 | 1259 (93) | 487 (95.1) | 1 | 782 (92.7) | 967 (94.6) | 1 |
| rs2746462 | GT | 120 | 26.00±0.41 | 87.81±1.27 | 95(7) | 25 (4.9) | 0.67(0.42-1.07) | 62(7.3) | 55 (5.4) | 0.69(0.46-1.03) |
| HWE:0.26 | p-valor | | 0.0026 | 0.0210 | | | 0.086 | | | 0.072 |
| NDUFS1 | CC | 1764 | 27.26±0.12 | 90.58±0.34 | 1293 (93.8) | 471 (90.9) | 1 | 791 (92.4) | 968 (93.3) | 1 |
| rs77655991 | CT | 128 | 27.96±0.48 | 92.03±1.38 | 84 (6.1) | 44 (8.5) | 1.50(1.01-2.24) | 64 (7.5) | 66 (6.4) | 0.85(0.58-1.25) |
| HWE:0.31 | TT | 4 | 28.58±1.93 | 96.25±6.68 | 1 (0.1) | 3 (0.6) | 1.58(1.09-2.29) | 1 (0.1) | 3 (0.3) | 0.88(0.61-1.26) |
| | p-valor | | 0.13 | 0.17 | | | 0.0180 | | | 0.48 |

Tabla 46. Asociaciones con OB de los SNPs seleccionados de la fase 2. La primera fila corresponde a los datos de la población total. OB C.: Obesidad Central.

5.3.3.2 Asociaciones con DMT2 de los polimorfismos seleccionados, en la población total de la fase 1

Al analizar las variantes identificadas en la secuenciación de los genes asociados con DMT2 y OB, no se han hallado asociaciones significativas con DMT2 en la población total, ni en las poblaciones femenina, masculina y obesa. Solo se ha visto una relación del genotipo TT del rs77655991 con mayor riesgo de DMT2 en los individuos no obesos. Los resultados de esta asociación se detallan en la tabla 47.

| Gen dbSNP | Genotipo | No T2DM (n(%)) | T2DM (n(%)) | Riesgo de T2DM OR(95% CI) |
|---------------|----------|----------------|-------------|---------------------------|
| | | | | p-valor |
| NDUFS1 | CC-CT | 2396 (99.9) | 167 (98.8) | 1.00 |
| rs77655991 | TT | 2 (0.1) | 2 (1.2) | 13.04 (1.64-103.80) |
| HWE:1 | p-valor | | | 0.026 |

Tabla 47. Asociación con la DMT2 en la población no obesa.

5.3.3.3 Asociaciones con HTA de los polimorfismos seleccionados, en la población de la fase 1

Al estudiar el riesgo de HTA, se ha encontrado una relación con el alelo C del rs1801318 y mayor riesgo, en la población total, en la femenina, y en la obesa. En mujeres, se ha visto que la asociación es más significativa, por lo que podría haber un efecto del género sobre este SNP y su relación con HTA. Sin embargo, es en los obesos en donde se ha encontrado la asociación más significativa, lo que podría significar que la obesidad es un factor importante en la relación de este SNP con el riesgo de HTA. Los resultados detallados se indican en la tabla 48. En las poblaciones masculina y no obesa, no hemos encontrado ninguna asociación relevante.

Con respecto al estudio de PAS y PAD, no se han hallado asociaciones significativas en ninguna de las poblaciones analizadas, a excepción de la **población femenina**, en donde se ha visto una con el genotipo CC del rs1801318 y menores valores de PAS. Los datos de la asociación se detallan en la tabla 49.

| Gen dbSNP | Genotipo | No HTA (n(%)) | HTA (n(%)) | Riesgo de HTA OR (95% CI) p-valor |
|-----------------------------------------|----------|---------------|------------|--------------------------------------|
| NDUFS1 rs1801318 HWE:0.032 | TT | 677 (43.3) | 486 (39.4) | 1.00 |
| | TC-CC | 887 (56.7) | 748 (60.6) | 1.28 (1.08-1.52) |
| | p-valor | | | 0.0044 |
| MUJERES | | | | |
| NDUFS1 rs1801318 HWE:0.099 | TT | 380 (43.3) | 257 (39) | 1.00 |
| | TC-CC | 498 (56.7) | 402 (61) | 1.45 (1.14-1.85) |
| | p-valor | | | 0.0025 |
| OBESOS | | | | |
| NDUFS1 rs1801318 HWE:0.87 | TT | 116 (51.6) | 177 (37.4) | 1.00 |
| | TC-CC | 109 (48.4) | 296 (62.6) | 1.85 (1.32-2.59) |
| | p-valor | | | 0.0004 |

Tabla 48. Asociaciones con la HTA en las poblaciones donde hubo algún hallazgo.

| Gen dbSNP | Genotipo | Número de pacientes (N) | PAS (Media en mg/dl(SD)) | Diferencias (95% IC) |
|------------------------------------------|----------|-------------------------|--------------------------|------------------------------|
| NDUFS1 rs1801318 HWE: 0.069 | TT-TC | 757 | 125.77 (0.68) | 0.00 |
| | CC | 77 | 120.57 (2.08) | -5.20 (-8.84 - -1.55) |
| | p-valor | | | 0.0054 |

Tabla 49. Asociación con PAS en la población femenina.

5.4 Fase 4: Estudios de replicación de las asociaciones halladas

Para confirmar las asociaciones halladas en 17 genes de la CMR, en la fase 1 y 3, en los estudios Horteiga, Pizarra y Segovia, se han analizado los 26 SNPs más significativos, en la población Di@bet.es, mediante secuenciación de nueva generación. Del total de polimorfismos, 23 correspondían a la fase 1, y 3 a la fase 3. Una vez finalizada la secuenciación masiva, se ha procedido al análisis bioinformático de los resultados. En primer lugar, se ha comprobado que todos los parámetros específicos de la secuenciación sean correctos según las especificaciones de Illumina. Después de obtener los archivos FastQ, se ha realizado el control de calidad de los datos obtenidos a partir de dichos archivos. Por último, se han calculado las coberturas de los amplicones secuenciados en todas las muestras.

Finalmente, se ha realizado el protocolo (pipeline) bioinformático para obtener los genotipos de los polimorfismos de interés, y continuar con los estudios de asociación con OB, DMT2 y HTA en la población.

Se tomaron como criterios de aceptación de los genotipos obtenidos en la secuenciación, un número mínimo de lecturas para cada SNP y para cada muestra, de 20x, y un Qscore, mayor a 30 (Q30) para cada base. Se incluyeron en estudios posteriores los SNPs con un porcentaje mayor al 70% de las muestras secuenciadas que cumplen con estos criterios, por lo que se ha excluido 1 SNP que estaba por debajo de ese valor (rs803224), para el estudio de asociación. Al evaluar el equilibrio de Hardy-Weinberg (p -valor > 0.002 para cumplir este equilibrio), se ha excluido el polimorfismo rs1801318, ya que ha sido el único que no lo ha cumplido.

5.4.1 Características generales de la población analizada

La población Di@bet.es utilizada para la réplica de los resultados obtenidos en la primera y tercera fase, posee las características generales que se detallan en la tabla 50, en la columna correspondiente. En esta tabla también se muestran los valores medios de los parámetros que definen a la población conjunta de los estudios Pizarra, Horteiga y Segovia, para poder comparar los valores de las 2 poblaciones. Es

importante recalcar que todas las variables que caracterizan a las poblaciones han mostrado diferencias significativas entre sus medias o sus porcentajes, excepto para la edad y la cintura.

| Variables | Población Di@bet.es (n=5544) | Población total fase 1 y 3 (n=3731) | Comparación entre las poblaciones (p-valor) |
|--------------------------|---------------------------------|----------------------------------------|------------------------------------------------------|
| Edad (años) | 51.0 (17.6) | 51.4 (15.9) | 0.259 |
| Género (H(%) / M(%)) | 2308(41.6) / 3236 (58.4) | 1681(45.1) / 2043(54.9) | 0.001 |
| IMC (kg/m ²) | 28.1 (5.2) | 27.4 (4.6) | < 0.001 |
| Cintura (cm) | 94.0 (14.2) | 93.6 (13.1) | 0.204 |
| Glucemia (mg/dl) | 99.08 (29.456) | 95.69 (24.132) | < 0.001 |
| PAS (mm Hg) | 131.7 (20.3) | 129.6 (20.3) | < 0.001 |
| PAD (mm Hg) | 77.0 (11.1) | 79.7 (11.0) | < 0.001 |
| OB (N(%)) | 1546 (30.5) | 912 (25.3) | < 0.001 |
| OB Central (N/(%)) | 3640 (48.0) | 1614 (44.6) | 0.002 |
| DMT2 (N/(%)) | 790 (15.4) | 434 (12.0) | < 0.001 |
| HTA (N/(%)) | 2430 (46.6) | 1798 (49.5) | 0.006 |

Tabla 50. Comparación de la población total formada por los estudios Horteiga, Pizarra y Segovia, y la población Di@bet.es. Los valores de esta tabla se expresan como la media (SD) excepto para las variables sexo, OB, OB Central, DMT2 y HTA, en las que se indica el número de pacientes (%).

5.4.2 Control de calidad de la secuenciación

-Comprobación de las especificaciones de los datos de secuenciación en el Miseq

Un resumen de los valores obtenidos para los parámetros utilizados en la evaluación de la secuenciación, para cada tipo de lectura (read) por separado, se muestra en la figura 40.

| Read 1 | | | | | | | | | | | |
|------------|-------|------------------------------|----------------|------------------|-----------|--------------|----------|-----------|------------------|---------------|----------------|
| Lane | Tiles | Density (K/mm ²) | Cluster PF (%) | Phas/Prephas (%) | Reads (M) | Reads PF (M) | % >= Q30 | Yield (G) | Cycles Err Rated | Aligned (%) | Error Rate (%) |
| 1 | 28 | 795 +/- 9 | 86,05 +/- 2,33 | 0,114 / 0,088 | 14,62 | 12,58 | 93,5 | 1,9 | 150 | 7,34 +/- 0,35 | 0,52 +/- 0,13 |
| Read 2 (I) | | | | | | | | | | | |
| Lane | Tiles | Density (K/mm ²) | Cluster PF (%) | Phas/Prephas (%) | Reads (M) | Reads PF (M) | % >= Q30 | Yield (G) | Cycles Err Rated | Aligned (%) | Error Rate (%) |
| 1 | 28 | 795 +/- 9 | 86,05 +/- 2,33 | 0,000 / 0,000 | 14,62 | 12,58 | 84,1 | 0,1 | 0 | 0,00 +/- 0,00 | 0,00 +/- 0,00 |
| Read 3 (I) | | | | | | | | | | | |
| Lane | Tiles | Density (K/mm ²) | Cluster PF (%) | Phas/Prephas (%) | Reads (M) | Reads PF (M) | % >= Q30 | Yield (G) | Cycles Err Rated | Aligned (%) | Error Rate (%) |
| 1 | 28 | 795 +/- 9 | 86,05 +/- 2,33 | 0,000 / 0,000 | 14,62 | 12,58 | 88,3 | 0,1 | 0 | 0,00 +/- 0,00 | 0,00 +/- 0,00 |
| Read 4 | | | | | | | | | | | |
| Lane | Tiles | Density (K/mm ²) | Cluster PF (%) | Phas/Prephas (%) | Reads (M) | Reads PF (M) | % >= Q30 | Yield (G) | Cycles Err Rated | Aligned (%) | Error Rate (%) |
| 1 | 28 | 795 +/- 9 | 86,05 +/- 2,33 | 0,074 / 0,129 | 14,62 | 12,58 | 86,2 | 1,9 | 150 | 7,15 +/- 0,29 | 0,71 +/- 0,33 |

Figura 40. Resumen de los parámetros necesarios para evaluar el correcto funcionamiento de la secuenciación en el equipo Miseq de Illumina.

Al comparar los valores de nuestra carrera con los definidos por Illumina, la cantidad de clusters que han pasado el filtro de calidad (684 K/mm²) está por debajo de lo esperado. A pesar de esto, las lecturas que han pasado el filtro de calidad (14,62 M x2) están dentro del rango establecido, y el porcentaje de bases con un Qscore mayor a Q30 es mayor al 80%. Por otro lado, el rendimiento total obtenido (4 Gb) no ha llegado al mínimo esperado, aunque su diferencia es de 0.5.

El porcentaje global de alineamiento del Phix es de 7.245, y en realidad esperábamos un 10 %. El porcentaje medio de error de la secuenciación, en base al Phix, fue de 0.615, es decir, normal.

-Evaluación de los archivos FastQ

Para evaluar la calidad de las secuencias obtenidas en la secuenciación realizada, se ha utilizado el programa FastQC. Se han seleccionado los controles específicos para el modo de secuenciación por amplicones. Sus valores para las lecturas 1 y 2 se resumen en la tabla 51.

| Parámetro | Muestras que pasan | Muestras con alerta | Muestras que fallan |
|-----------------------------|--------------------|---------------------|---------------------|
| Per base N content | 9603 | 117 | 0 |
| Per base sequence quality | 9620 | 24 | 76 |
| Per sequence quality scores | 9711 | 1 | 8 |

Tabla 51. Comportamiento de las muestras según los distintos controles evaluados en los FastQ.

A continuación se muestran como ejemplo, los gráficos de los parámetros obtenidos para una muestra en particular (Muestra 41002 del Estudio Di@bet.es) correspondientes a la lectura 1, después del análisis en FastQC:

-Contenido de N por base: se ha comprobado que el porcentaje de bases N en cada base secuenciada era igual a cero. En la gráfica de la figura 41 se puede ver el valor se mantiene a lo largo de toda la lectura

-Calidad de secuencia por base: se ha observado que el Qscore medio era mayor a 30 a lo largo de toda la lectura, indicando que los datos obtenidos tras la secuenciación tienen la calidad adecuada (figura 42).

-Scores de calidad por secuencia: se ha podido ver que el Qscore tiene una distribución correcta en todas las secuencias, se comportan de manera homogénea (figura 43).

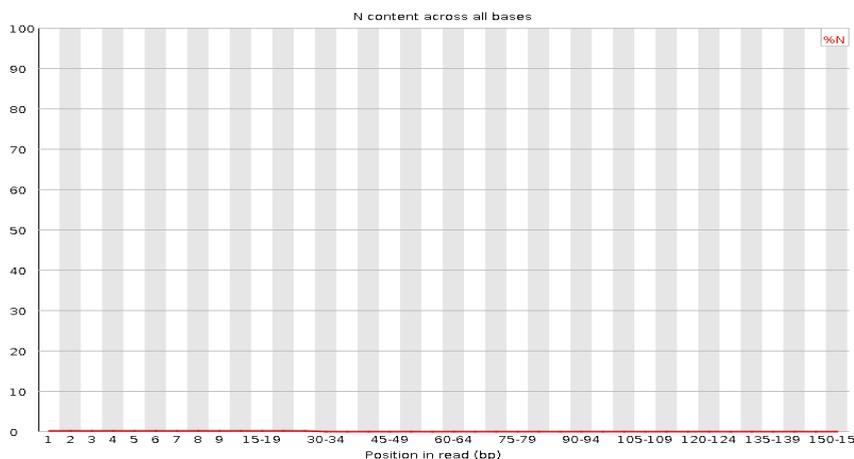


Figura 41. Contenido de N por base para la muestra 41002 del Estudio Diab@t.es

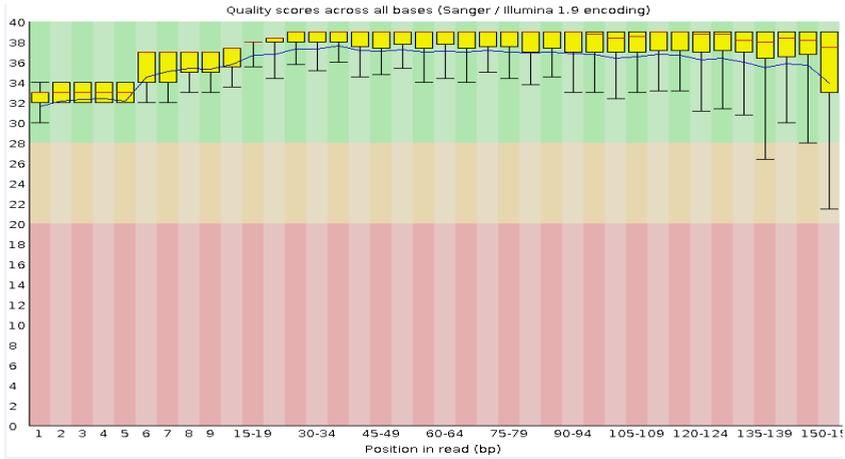


Figura 42. Calidad de la secuencia por base para la muestra 41002 del Estudio Di@bet.es. La línea roja representa la mediana y la línea azul la calidad media de secuenciación. -Scores de calidad por secuencia.

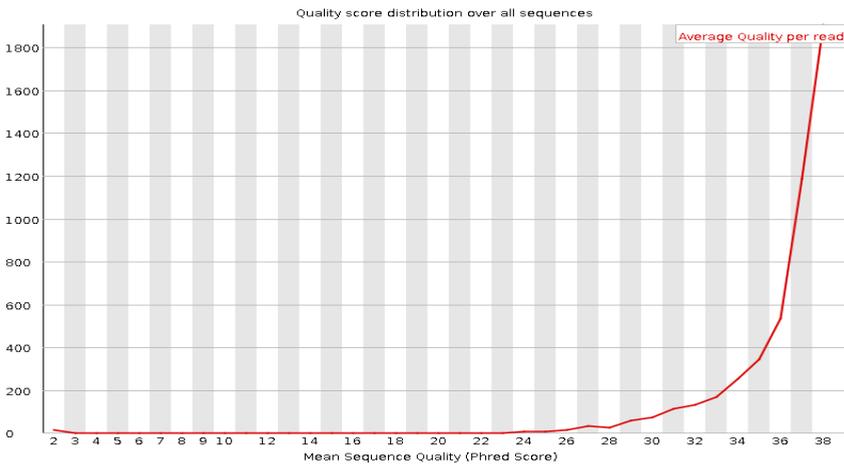


Figura 43. Scores de calidad por secuencia para la muestra 41002 del Estudio Di@bet.es.

-Distribución del número de lecturas según los fastQ (Figura 44): se han realizado los cálculos de distribución del número de lecturas obtenidas para todas las muestras, según los valores obtenidos en los fastQ, y se han representado en dos formas de gráficos de distribución, en histograma (figura A) y en gráfico de densidad (figura B).

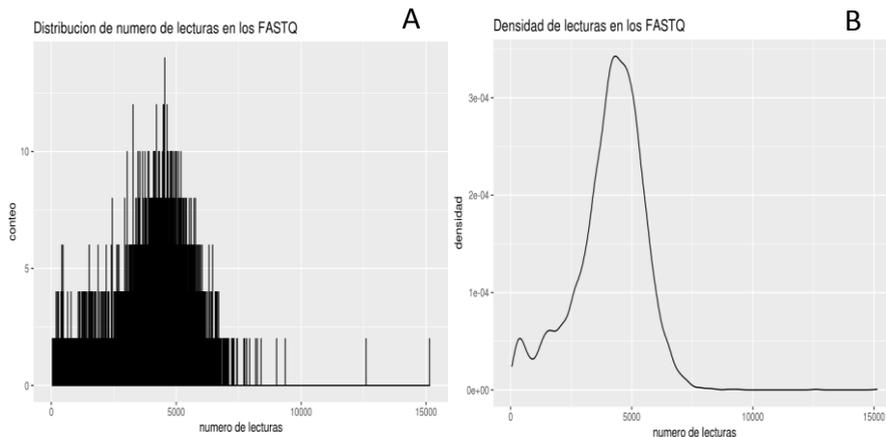


Figura 44. Distribución del número de lecturas según los fastQ para el total de muestras analizadas, graficada en Histograma (A) y en densidad de lecturas (B).

El número medio de lecturas en las muestras secuenciadas ha sido de 4097, con un mínimo de 46 y un máximo de 15140.

Por todo lo expuesto en este apartado, los resultados de todos los parámetros de control de calidad, nos han permitido concluir, que las secuencias obtenidas han mostrado calidades correctas para este tipo de secuenciaciones.

5.4.3 Cobertura de los amplicones secuenciados

Después de realizar el control de calidad de la secuenciación, se han analizado los datos de cobertura de las secuencias obtenidas, para los 26 amplicones estudiados. Para todos los SNPs, y todas las muestras, la cobertura obtenida ha sido mayor a 20x, por lo que se han obtenido unos valores adecuados para el uso de los genotipos obtenidos en la secuenciación, en los análisis de asociación posteriores. Los valores medios de cobertura para cada amplicón se detallan en la figura del anexo IV.

5.4.4 Asociaciones de los polimorfismos de la CMR y la OB en la población Di@bet.es

En la tabla 52 se detallan todas las asociaciones halladas al estudiar los 26 polimorfismos seleccionados de la fase 1 y 3, en la población Di@bet.es, con todos los parámetros relacionados con OB. Los SNPs rs7255180 y rs535277 se han

asociado con el riesgo de OB y con el parámetro que define la obesidad, el IMC. Por otro lado, el alelo G del rs535277 se ha asociado con valores bajos de cintura y de riesgo de obesidad central.

5.4.4.1 Resultados de asociaciones de SNPs de CMR y OB en subpoblaciones

5.4.4.1.1 Resultados de las asociaciones de CMR Y OB según género

Se han obtenido las asociaciones con OB e IMC detalladas en la tabla 53, para las poblaciones masculinas y femeninas. Los polimorfismos cuyas asociaciones se han encontrado con IMC y OB de forma simultánea, en este caso, los del gen NDUFS4 en mujeres, y los de los genes NDUFS2, COX5B, COX7A1 y 2 en hombres, podrían estar afectando al riesgo de OB a través de la modificación del IMC.

| Gen | Genotipos | N | IMC Kg/m2 (SD) | Cintura cm (SD) | No OB (n%) | OB (n%) | OR/95% CI | No OB Central (n%) | OB Central (n%) | OR/95% CI |
|---------------------------------|-----------|------|-------------------|--------------------|-------------|-------------|------------------------|-----------------------|--------------------|------------------------|
| COX5A rs1133322 HWE:0.27 | AA | 1247 | 27.8(0.14) | 93.47 (0.38) | 896 (28.5) | 351 (25.6) | 1 | 669 (28.4) | 580 (26.8) | 1 |
| | AG | 991 | 28.0(0.11) | 93.99 (0.3) | 1569 (49.9) | 708 (51.6) | 1.15(0.99-1.35) | 1185 (50.2) | 1096 (50.6) | 1.05(0.91-1.22) |
| | GG | | 28.2(0.16) | 94.62 (0.44) | 679 (21.6) | 312 (22.8) | 1.18(0.98-1.42) | 505 (21.4) | 489 (22.6) | 1.14(0.95-1.36) |
| | p-valor | | 0.044 | 0.026 | | | 0.071 | | | 0.17 |
| COX7A1 rs7255180 HWE:0.63 | CC | 3951 | 27.9(0.08) | 93.75 (0.22) | 2774 (88.3) | 1177 (85.7) | 1 | 2085 (88.2) | 1872 (86.6) | 1 |
| | CT-TT | 565 | 28.6(0.22) | 95.43 (0.6) | 369 (11.7) | 196 (14.3) | 1.23(1.02-1.49) | 278 (11.8) | 289 (13.4) | 1.15(0.95-1.39) |
| | p-valor | | 0.0069 | 0.028 | | | 0.033 | | | 0.16 |
| NDUFS4 rs535277 HWE:0.6 | AA | 2409 | 28.2(0.11) | 94.66 (0.29) | 1636 (67.4) | 773 (70.6) | 1 | 1204 (67.1) | 1212 (69.9) | 1 |
| | AG-GG | 1114 | 27.8(0.15) | 93.8 (0.41) | 792 (32.6) | 322 (29.4) | 0.84(0.72-0.99) | 591 (32.9) | 523 (30.1) | 0.85(0.73-1.00) |
| | p-valor | | 0.028 | 0.016 | | | 0.033 | | | 0.044 |
| COX5B rs1470625 HWE:0.49 | AA | 2286 | 27.9(0.1) | 93.71 (0.29) | 1625 (52.8) | 657 (48.9) | 1 | 1206 (52.4) | 1080 (50.8) | 1 |
| | CA-CC | 2144 | 28.1(0.11) | 94.39 (0.32) | 1453 (47.2) | 687 (51.1) | 1.18(1.04-1.35) | 1096 (47.6) | 1048 (49.2) | 1.09(0.96-1.24) |
| | p-valor | | 0.14 | 0.04 | | | 0.012 | | | 0.18 |
| NDUFS2 rs1136224 HWE:0.83 | AA-AG | 4058 | 28.3(0.08) | 94.05 (0.22) | 2821 (97.7) | 1237 (98.1) | 1 | 2106 (97.4) | 1960 (98.3) | 1 |
| | GG | 92 | 27.2(0.47) | 91.22 (1.43) | 68 (2.4) | 24 (1.9) | 0.83(0.51-1.34) | 57 (2.6) | 34 (1.7) | 0.61(0.38-0.97) |
| | p-valor | | 0.23 | 0.19 | | | 0.44 | | | 0.033 |

Tabla 52. Asociaciones de los polimorfismos de la CMR y todos los parámetros relacionados con la OB en la población Di@bet.es.

En el caso de los resultados obtenidos para la Cintura y la OB Central, todas las variantes halladas aquí, a excepción del rs17624157, se han asociado también con IMC y/o OB, en las mismas subpoblaciones y con las mismas tendencias. Por otro lado, sólo uno de los SNPs podría estar afectando al riesgo de padecer OB Central por su posible efecto sobre la Cintura: el rs7255180 en hombres. Los detalles de esta asociación, junto con el resto de las halladas en las poblaciones masculinas y femeninas, se indican en la tabla 54.

Al evaluar el comportamiento de las asociaciones halladas en estas subpoblaciones respecto a la población total, se ha visto que en mujeres, la asociación del rs535277 ha sido más significativa para IMC, OB y Cintura. En el caso de los hombres, las asociaciones con IMC y Cintura del SNP rs1470625 han mejorado, no así las que se habían hallado con OB. Para el rs7255180, se ha encontrado una asociación menos significativa en IMC, pero por el contrario, unas más significativas con Cintura y OB Central.

| Población Diab@Les | | N | | IMC (Kg/m2) | | No Obesidad (n%) | | Obesidad (n%) | | OR(95% CI) | |
|----------------------------------|---------|------|-------|--------------|--------------|------------------|-------------|---------------|------------|-----------------|-----------------|
| Gen/SNP | G | H | M | H | M | H | M | H | M | H | M |
| NDUFS4 rs535277 HWE:0.81 | AA | 1012 | 1397 | 28.58 (0.14) | 27.91 (0.15) | 682 (67.6) | 954 (67.2) | 330 (68.2) | 443 (72.5) | 1 | 1 |
| | AG-GG | 481 | 633 | 28.55 (0.2) | 27.3 (0.22) | 327 (32.4) | 465 (32.8) | 154 (31.8) | 168 (27.5) | 0.98(0.77-1.23) | 0.73(0.59-0.91) |
| | p-valor | | | 0.93 | 0.0046 | | | | | 0.84 | 0.0045 |
| COX5A rs1133322 HWE:1 | AA | 546 | 701 | 28.22 (0.19) | 27.5 (0.19) | 388 (29.2) | 508 (28) | 158 (25.6) | 193 (25.6) | 1.00 | 1.00 |
| | AG | 963 | 1314 | 28.53 (0.14) | 27.68 (0.16) | 651 (49.1) | 918 (50.5) | 312 (50.6) | 396 (52.5) | 1.17(0.93-1.47) | 1.15(0.93-1.42) |
| | GG | 435 | 556 | 28.86 (0.21) | 27.72 (0.23) | 288 (21.7) | 391 (21.5) | 147(23.8) | 165 (21.9) | 1.24(0.94-1.63) | 1.16(0.90-1.50) |
| p-valor | | | 0.03 | 0.26 | | | | | 0.12 | 0.23 | |
| COX5B rs1470625 HWE:1 | AA | 983 | 1299 | 28.3 (0.14) | 27.65 (0.15) | 692 (53.4) | 933 (52.4) | 291 (48.1) | 366 (49.5) | 1 | 1 |
| | CA-CC | 919 | 1221 | 28.73 (0.15) | 27.68 (0.17) | 605 (46.6) | 848 (47.6) | 314 (51.9) | 373(50.5) | 1.25(1.03-1.52) | 1.12(0.93-1.33) |
| | p-valor | | | 0.023 | 0.96 | | | | | 0.024 | 0.23 |
| COX7A2 rs683943 HWE:1 | GG-CG | 1857 | 2467 | 28.51 (0.1) | 27.64 (0.11) | 1269 (100) | 1745 (99.8) | 588 (99.3) | 722 (99.9) | 1 | 1 |
| | CC | 4 | 5 | 34.17 (1.47) | 26.13 (1.4) | 0(0) | 4(0.2) | 4(0.7) | 1(0.1) | NA(0.00-NA) | 0.60(0.06-5.72) |
| | p-valor | | | 0.009 | 0.51 | | | | | 0.0023 | 0.64 |
| NDUFS2 rs4656993 HWE:0.92 | GG | 688 | 945 | 28.23 (0.16) | 27.58 (0.19) | 485 (37.7) | 665 (37.4) | 203 (34.1) | 280 (38) | 1 | 1 |
| | AG | 899 | 1223 | 28.54 (0.15) | 27.71 (0.16) | 615 (47.8) | 862 (48.5) | 284 (47.7) | 361 (49) | 1.11(0.89-1.38) | 0.95(0.78-1.16) |
| | AA | 294 | 345 | 28.87 (0.27) | 27.51 (0.29) | 18 6(14.5) | 249 (14) | 108 (18.1) | 96(13) | 1.38(1.03-1.84) | 0.94(0.71-1.25) |
| p-valor | | | 0.035 | 0.94 | | | | | 0.037 | 0.61 | |
| COX7A1 rs7255180 HWE:1 | CC | 1688 | 2263 | 28.41 (0.11) | 27.58 (0.12) | 1166 (88.2) | 1608 (88.3) | 522 (84.6) | 655 (86.6) | 1 | 1 |
| | CT-TT | 251 | 314 | 29.22 (0.28) | 28.15 (0.32) | 156 (11.8) | 213 (11.7) | 95 (15.4) | 101 (13.4) | 1.34(1.02-1.77) | 1.15(0.88-1.50) |
| | p-valor | | | 0.0091 | 0.13 | | | | | 0.038 | 0.3 |
| COX6A1 rs17431357 HWE:0.53 | CC | 1745 | 2368 | 28.58 (0.11) | 27.66 (0.11) | 1174 (86.6) | 167 7(89) | 571 (91.1) | 691 (88.1) | 1 | 1 |
| | CT | 229 | 289 | 27.99 (0.28) | 27.71 (0.34) | 173 (12.8) | 200 (10.6) | 56 (8.9) | 89 (11.3) | 0.68(0.49-0.93) | 1.02(0.77-1.34) |
| | TT | 8 | 11 | 27.23 (0.7) | 26.57 (1.41) | 8(0.6) | 7(0.4) | 0(0) | 4(0.5) | 0.00(0.00-NA) | 1.30(0.35-4.77) |
| p-valor | | | 0.053 | 0.47 | | | | | 0.0024 | 0.79 | |
| NDUFS1 rs3770989 HWE:0.26 | TT | 1380 | 1906 | 28.59 (0.12) | 27.72 (0.13) | 930 (91.8) | 1338 (93.4) | 450 (94.7) | 568 (93.6) | 1 | 1 |
| | TC | 108 | 133 | 28.25 (0.41) | 27.79 (0.46) | 83(8.2) | 94 (6.6) | 25 (5.3) | 39 (6.4) | 0.62(0.39-0.98) | 0.98(0.65-1.46) |
| | p-valor | | | 0.42 | 0.87 | | | | | 0.035 | 0.91 |

Tabla 53. Asociaciones obtenidas con la OB en las poblaciones femenina y masculina de Di@bet.es. G: genotipo; H: hombres; M: mujeres.

| Población Di@bet.es | | N° de pacientes | | Cintura (cm) | | No obesidad central (n/%) | | Obesidad central (n/%) | | OR/95% CI | |
|---------------------|-----------|-----------------|------|---------------|---------------|---------------------------|------------|------------------------|------------|------------------------|-----------------|
| dbSNP | Genotipos | H | M | H | M | H | M | H | M | H | M |
| rs1470625 | AA | 987 | 1299 | 98.23(0.37) | 90.27(0.39) | 625(53.2) | 581(51.5) | 362(49.4) | 718(51.5) | 1 | 1 |
| | CA-CC | 921 | 1223 | 99.26(0.4) | 90.73(0.44) | 550(46.8) | 546(48.5) | 371(50.6) | 677(48.5) | 1.21(1.00-1.47) | 1.00(0.84-1.19) |
| | p-valor | | | 0.018 | 0.46 | | | | | 0.052 | 0.99 |
| rs4656993 | GG | 690 | 946 | 97.83(0.44) | 90.49(0.5) | 449(38.3) | 429(38.1) | 241(33.7) | 517(37.2) | 1 | 1 |
| | AG | 903 | 1227 | 98.9(0.39) | 90.63(0.41) | 548(46.8) | 540(48) | 355(49.6) | 687(49.4) | 1.23(0.99-1.53) | 1.00(0.83-1.20) |
| | AA | 294 | 343 | 99.43(0.75) | 89.71(0.77) | 174(14.9) | 156(13.9) | 120(16.8) | 187(13.4) | 1.27(0.95-1.71) | 1.03(0.78-1.35) |
| HWE:1 | p-valor | | | 0.036 | 0.49 | | | | | 0.053 | 0.87 |
| rs7255180 | CC | 1693 | 2264 | 98.37(0.29) | 90.3(0.31) | 1066(88.5) | 1019(87.9) | 627(84.6) | 1245(87.7) | 1 | 1 |
| | CT-TT | 252 | 315 | 100.71(0.7) | 91.2(0.82) | 138(11.5) | 140(12.1) | 114(15.4) | 175(12.3) | 1.39(1.05-1.83) | 0.97(0.75-1.26) |
| | p-valor | | | 0.0062 | 0.46 | | | | | 0.022 | 0.84 |
| rs17431357 | CC | 1748 | 2372 | 98.86(0.29) | 90.47(0.3) | 1063(86.5) | 1066(89.1) | 685(90.4) | 1306(88.5) | 1 | 1 |
| | CT | 231 | 288 | 97.33(0.72) | 90.73(0.88) | 159(12.9) | 125(10.4) | 72(9.5) | 163(11.1) | 0.71(0.53-0.97) | 0.97(0.74-1.28) |
| | TT | 8 | 11 | 98.88(3.53) | 91(4.95) | 7(0.6) | 5(0.4) | 1(0.1) | 163(11.1) | 0.19(0.02-1.61) | 0.87(0.23-3.37) |
| HWE:0.53 | p-valor | | | 0.13 | | 0.68 | | | | 0.0086 | 0.81 |
| rs17624157 | GG-GA | 1908 | 2563 | 98.61(0.27) | 90.5(0.29) | 1186(98.1) | 1149(98) | 722(96.4) | 1414(98.1) | 1 | 1 |
| | AA | 50 | 51 | 101.82(1.7) | 87.94(1.61) | 23(1.9) | 24(2) | 27(3.6) | 27(1.9) | 1.99(1.10-3.58) | 0.94(0.51-1.75) |
| | p-valor | | | 0.062 | 0.21 | | | | | 0.023 | 0.85 |
| rs535277 | AA | 1016 | 1400 | 90.5(0.29) | 91.39(0.4) | 611(66.8) | 593(67.4) | 405(69.5) | 807(70) | 1 | 1 |
| | AG-GG | 482 | 632 | 98.6(0.53) | 90.15(0.56) | 304(33.2) | 287(32.6) | 178(30.5) | 345(29.9) | 0.88(0.70-1.12) | 0.82(0.66-1.01) |
| | p-valor | | | 0.43 | 0.0098 | | | | | 0.3 | 0.062 |

Tabla 54. Resultados de las asociaciones con Cintura y OB Central en hombres y mujeres. H: hombres; M: mujeres.

5.4.5 Asociaciones de los polimorfismos de la CMR y la DMT2 en Di@bet.es

Las asociaciones de las variantes a replicar, con la DMT2, se detallan en la tabla 55. Tres de los SNPs asociados con diabetes, se han asociado al mismo tiempo con OB (rs1470625, rs7255180 y rs1136224).

Los detalles de las 2 asociaciones halladas con glucemia se indican en la tabla 56.

| dbSNP | Genotipo | No DMT2 (n(%)) | DMT2 (n(%)) | Riesgo de DMT2 OR(95% CI) p-valor |
|--------------------------------|---------------------------|------------------------------------------|----------------------------------------|-----------------------------------------------------------------|
| rs1470625 HWE:0.12 | AA CA-CC p-valor | 1996 (52.3) 1821 (47.7) | 285 (47.2) 319 (52.8) | 1 1.26 (1.04-1.52) 0.017 |
| rs7255180 HWE:0.63 | CC CT-TT p-valor | 3403 (87.2) 498 (12.8) | 547 (89.1) 67 (10.9) | 1 0.73 (0.55-0.98) 0.035 |
| rs4600063 HWE:0.082 | AA-AG GG p-valor | 3891 (99.4) 23 (0.6) | 622 (100) 0 (0) | 1 0.0 (0.00-NA) 0.024 |
| rs683943 HWE:0.36 | GG-CG CC p-valor | 3740 (99.9%) 5 (0.1%) | 583 (99.3%) 4 (0.7%) | 1 7.27 (1.65-31.98) 0.013 |
| rs4806187 HWE: 0.011 | AA-AG GG p-valor | 2639 (77.8%) 751 (22.1%) | 413 (73.8%) 147 (26.2%) | 1 1.27 (1.01-1.59) 0.041 |
| rs1136224 HWE:0.5 | AA AG GG p-valor | 2599 (72.4%) 904 (25.2%) 87 (2.4%) | 413 (73.9%) 141 (25.2%) 5 (0.9%) | 1 1.07 (0.85-1.34) 0.36 (0.14-0.95) 0.05 |

Tabla 55. Resultados de las asociaciones obtenidas con DMT2 en Di@bet.es.

| SNP | Genotipos | N | Glucemia (Media en mg/dl(SD)) | Diferencias (95% IC) |
|------------------------|------------------------|--------------|----------------------------------|------------------------------------------------------|
| rs10504961 HWE:0.75 | CC CT-TT p-valor | 1178 2720 | 93.07 (0.55) 94.85 (0.39) | 0.00 1.44 (0.18 - 2.70) 0.025 |
| rs11205591 HWE:0.23 | CC-CG GG p-valor | 3175 188 | 94.06 (0.35) 90.52 (0.99) | 0.00 -3.13 (-5.84 - -0.41) 0.024 |

Tabla 56. Asociación obtenida para la glucemia en la población Di@bet.es.

5.4.5.1 Resultados de asociaciones de SNPs de CMR y en subpoblaciones

5.4.5.1.1 Resultados de las asociaciones de CMR Y DMT2 según género

Al analizar los resultados obtenidos con DMT2 en mujeres, se han encontrado asociaciones más significativas en mujeres, en comparación con la población total, para los SNPs rs1470625, rs4806187 y rs7255180. Con respecto a los hombres, los alelos G del polimorfismo rs1133322 y del rs9360898 han mostrado una asociación

con menor riesgo de DMT2. Los detalles de las asociaciones con DMT2 de ambas subpoblaciones se muestran en la tabla 57.

Al comparar las asociaciones de glucemia en las poblaciones de hombres y mujeres, con la muestra total, se ha visto que las relaciones de los SNPs rs10504961 y rs11205591 obtenidas en esta última, no estaban influenciadas por el género, debido a que ambos sexos se han comportado con la misma tendencia en sus valores de OR, a pesar de que sólo ha llegado a asociarse significativamente una de las subpoblaciones. Estos resultados, junto con las demás asociaciones con glucemia de las poblaciones masculina y femenina, se detallan en la tabla 58.

| Gen SNP Sexo | Genotipo | NO DMT2 | | DMT2 | | OR (95% CI) | |
|-------------------------------|----------|------------|------------|-----------|-----------|---------------------------|------------------------|
| | | (n/%) | | (n/%) | | p-valor | |
| | | Hombres | Mujeres | Hombres | Mujeres | Hombres | Mujeres |
| rs1133322 HWE:0.28 | AA | 434(27.2) | 642(27.9) | 112(32.4) | 59(21.8) | 1 | 1 |
| | AG | 804(50.3) | 1168(50.8) | 159(46) | 145(53.5) | 0.66(0.49-0.88) | 1.33(0.95-1.88) |
| | GG | 360(22.5) | 489(21.3) | 75(21.7) | 67(24.7) | 0.64(0.45-0.91) | 1.66(1.11-2.47) |
| | p-valor | | | | | 0.008 | 0.013 |
| rs1470625 HWE:0.83 | AA | 810(51.8) | 1186(52.7) | 173(51.3) | 112(42) | 1 | 1 |
| | CA-CC | 755(48.2) | 1066(47.3) | 164(48.7) | 155(58) | 1.03(0.80-1.33) | 1.56(1.18-2.06) |
| | p-valor | | | | | 0.8 | 0.0018 |
| rs7255180 HWE:1 | CC | 1388(87) | 2015(87.4) | 300(87.2) | 247(91.5) | 1 | 1 |
| | CT-TT | 207(13) | 291(12.6) | 44(12.8) | 23(8.5) | 0.86(0.59-1.25) | 0.59(0.37-0.95) |
| | p-valor | | | | | 0.43 | 0.022 |
| rs9360898 HWE:0.5 | TT | 983(61.4) | 1383(59.7) | 230(66.7) | 161(57.9) | 1.00 | 1.00 |
| | GT-GG | 619(38.6) | 932(40.3) | 115(33.3) | 117(42.1) | 0.70(0.54-0.92) | 1.12(0.85-1.47) |
| | p-valor | | | | | 0.0087 | 0.44 |
| rs1136224 HWE:1 | AA-AG | 1446(98) | 2057(97.3) | 308(99) | 246(99.2) | 1.00 | 1 |
| | GG | 30(2) | 57(2.7) | 3(1) | 2(0.8) | 0.86(0.63-1.17) | 0.28(0.06-1.24) |
| | p-valor | | | | | 0.32 | 0.048 |
| rs4656993 HWE:0.09 | GG-AG | 1297(83.8) | 1935(86.2) | 290(87.1) | 232(87.2) | 1 | 1 |
| | AA | 251(16.2) | 311(13.8) | 43(12.9) | 34(12.8) | 0.65(0.44-0.94) | 1.04(0.69-1.56) |
| | p-valor | | | | | 0.019 | 0.87 |
| rs4806187 HWE:0.015 | AA-AG | 1087(79.2) | 1552(77) | 218(71.2) | 195(76.8) | 1 | 1 |
| | GG | 286(20.8) | 465(23.1) | 88(28.8) | 59(23.2) | 0.99(0.70-1.39) | 1.60(1.18-2.18) |
| | p-valor | | | | | 0.95 | 0.0029 |
| rs683943 HWE:0.36 | GG-CG | 1534(99.9) | 2206(99.8) | 323(99.1) | 260(99.6) | 1.00 | 1.00 |
| | CC | 1(0.1) | 4(0.2) | 3(0.9) | 1(0.4) | 13.26(1.22-143.72) | 3.89(0.36-41.68) |
| | p-valor | | | | | 0.019 | 0.32 |

Tabla 57. Asociaciones con el riesgo de DMT2 en las poblaciones femenina y masculina del estudio Di@bet.es.

| SNP | Genotipo | Glucemia (Media en mg/dl(SD)) | | | | Diferencias (95%IC) | |
|------------------------|------------------|----------------------------------|------|------------|-------------|--------------------------------------------|-------------------------------------------|
| | | N | | Glucemia | | Diferencias (95%IC) | |
| Sexo | | H | M | H | M | H | M |
| rs10504961 HWE:0.28 | CC | 470 | 708 | 96.84(1.0) | 90.56(0.58) | 0.00 | 0.00 |
| | CT-TT p-valor | 1166 | 1554 | 98.96(0.7) | 91.76(0.41) | 1.50(-0.94-3.95) 0.23 | 1.34(0.06-2.63) 0.041 |
| rs1470625 HWE=1 | AA | 830 | 1153 | 98.35(0.8) | 90.73(0.42) | 0.00 | 0.00 |
| | CA-CC p-valor | 768 | 1049 | 98.01(0.8) | 91.99(0.51) | -0.50(-2.71-1.72) 0.66 | 1.45(0.27-2.63) 0.016 |
| rs39830 HWE=0.7 | TT | 557 | 797 | 100.3(1.2) | 90.71(0.56) | 0.00 | 0.00 |
| | CT | 794 | 1044 | 96.78(0.5) | 91.44(0.45) | -3.02(-5.50--0.55) | 1.05(-0.25-2.35) |
| | CC p-valor | 282 | 391 | 98.68(1.5) | 92.17(0.85) | -0.82(-4.09-2.45) 0.046 | 2.30(0.60-4.01) 0.0073 |
| rs535277 HWE=0.55 | AA | 856 | 1223 | 98.54(0.7) | 92.34(0.44) | 0.00 | 0.00 |
| | CA-CC p-valor | 405 | 555 | 98.6(1.25) | 90.33(0.63) | 0.00(-2.72-2.73) 1 | -1.57(-2.96--0.17) 0.028 |
| rs1133322 HWE=0.73 | AA | 467 | 624 | 99.87(1.3) | 90.67(0.59) | 0.00 | 0.00 |
| | AG | 808 | 1142 | 98.28(0.8) | 91.74(0.48) | -1.87(-4.46-0.72) | 1.11(-0.27-2.50) |
| | GG p-valor | 362 | 476 | 96.19(0.8) | 91.02(0.67) | -4.37(-7.49--1.24) 0.0065 | 0.62(-1.08-2.31) 0.4 |
| rs683943 HWE:0.36 | GG-CG | 1568 | 2156 | 98.01(0.5) | 91.33(0.33) | 0.00 | 0.00 |
| | CC p-valor | 3 | 4 | 128.5(15) | 84.38(3.69) | 26.66(1.18-52.15) 0.04 | -5.42(-19.30-8.47) 0.44 |
| rs1120591 HWE:0.23 | CC-CG | 1335 | 1840 | 98.21(0.6) | 91.04(0.36) | 0.00 | 0.00 |
| | GG p-valor | 89 | 99 | 92.35(1.3) | 88.87(1.41) | -4.97(-9.88--0.06) 0.047 | -1.43(-4.32-1.47) 0.34 |

Tabla 58. Asociaciones con la glucemia en las poblaciones femenina (M) y masculina (H) del estudio Di@bet.es.

5.4.5.1.2 Resultados de las asociaciones de CMR Y DMT2 en individuos obesos

En la población obesa, se ha hallado asociación con el genotipo CC del rs683943 y mayor riesgo de DMT2 (tabla 59), pero con menos significación que en la población total. Un comportamiento opuesto se encontró en el análisis de la glucemia, para el polimorfismo rs10504961, ya que se halló una asociación más significativa en obesos que en población general. El genotipo GG del rs4806187 se ha asociado con

valores mayores de glucemia en los obesos. Los resultados de los SNPs relacionados con glucemia se detallan en la tabla 60.

| Gen SNP | Genotipos | No DMT2 (n/%) | DMT2 (n/%) | OR (95 % CI) |
|---------------------------------------|-----------|---------------|------------|--------------------------|
| COX7A2 rs683943 HWE:0.71 | GG-CG | 1001 (99.8) | 308 (99) | 1.00 |
| | CC | 2 (0.2) | 3 (1) | 6.82 (1.01-45.81) |
| | p-valor | | | 0.048 |

Tabla 59. Asociación obtenida con la DMT2 en la población obesa.

| Gen SNP | Genotipos | N | Glucemia (Media en mg/dl(SD)) | Diferencias (95% IC) |
|-----------------------------------------|-----------|-----|-------------------------------|---------------------------|
| COX7A2 rs683943 HWE:0.71 | GG | 991 | 100.53 (0.65) | 0.00 |
| | CG | 94 | 105.04 (3.22) | 4.74 (0.26 - 9.22) |
| | CC | 4 | 117.85 (15.1) | 17.57 (-3.21 - 38.34) |
| | p-valor | | | 0.011 |
| COX6B1 rs4806187 HWE:0.015 | AA-AG | 770 | 100.48 (0.72) | 0.00 |
| | GG | 234 | 106 (2.02) | 5.46 (2.19 - 8.73) |
| | p-valor | | | 0.0011 |
| UQCRB rs10504961 HWE:0.28 | CC | 341 | 98.61 (0.99) | 0.00 |
| | CT-TT | 800 | 102.4 (0.84) | 3.48 (0.71 - 6.25) |
| | p-valor | | | 0.014 |

Tabla 60. Asociaciones halladas con la glucemia en la población obesa.

5.4.5.1.3 Resultados de las asociaciones de CMR Y DMT2 en individuos no obesos

Al analizar las asociaciones de los polimorfismos de la CMR con DMT2 en población no obesa, no se ha obtenido ningún hallazgo. Sin embargo, se han encontrado tres asociaciones con glucemia, aunque no muy significativas. Los resultados se detallan en la tabla 61.

| Gen SNP | Genotipos | N | Glucemia (Media en mg/dl(SD)) | Diferencias (95% IC) |
|-------------------------------------------------|-----------|------|-------------------------------|------------------------------|
| COX7C rs2410718 HWE:0.17 | GG | 2553 | 91.57 (0.36) | 0.00 |
| | AG | 248 | 89.49 (0.94) | -1.82 (-4.01 - 0.37) |
| | AA | 5 | 107.86 (16.91) | 15.98 (1.28 - 30.68) |
| | p-valor | | | 0.026 |
| NDUFS5 rs11205591 HWE:0.032 | CC | 1281 | 91.81 (0.57) | 0.00 |
| | CG | 959 | 90.44 (0.47) | -1.15 (-2.57 - 0.27) |
| | GG | 143 | 88.7 (1.08) | -1.27 (-2.39 - -0.15) |
| | p-valor | | | 0.026 |
| NDUFS2 rs4656993 HWE:0.53 | GG-AG | 2292 | 91.58 (0.36) | 0.00 |
| | AA | 387 | 89.56 (0.72) | -1.82 (-3.54 - -0.11) |
| | p-valor | | | 0.037 |
| | | | | |

Tabla 61. Asociaciones halladas con la glucemia en la población no obesa.

5.4.6 Asociaciones de los polimorfismos de la CMR y la HTA en la población Di@bet.es

Se han hallado dos polimorfismos asociados a HTA en la muestra total analizada. Por un lado, el alelo C del rs1133322 se ha asociado con menor riesgo de enfermedad, mientras que el alelo G del rs9360898 se ha relacionado con mayor riesgo. Los detalles se muestran en la tabla 62. Se encontró también una asociación muy significativa del rs1133322 y PAD (tabla 63), lo que podría indicar que este SNP afectaría al riesgo de HTA, modificando a la presión diastólica.

| Gen dbSNP | Genotipo | No HTA (n(%)) | HTA (n(%)) | Riesgo de HTA OR(95% CI) p-valor |
|-----------------------------------------------|----------|---------------|-------------|-------------------------------------|
| COX5A rs1133322 HWE:0.32 | TT | 663 (26.6) | 579 (28.8) | 1.00 |
| | CT-CC | 1833 (73.4) | 1433 (71.2) | 0.83 (0.70-0.97) |
| | p-valor | | | 0.02 |
| COX7A2 rs9360898 HWE:0.17 | TT | 1551 (61.7) | 1203 (59.5) | 1.00 |
| | GT-GG | 963 (38.3) | 818 (40.5) | 1.17 (1.01-1.36) |
| | p-valor | | | 0.038 |

Tabla 62. Asociaciones halladas con la HTA en la población de Di@bet.es.

| Gen dbSNP | Genotipo | Número de pacientes (N) | PAD (Media en mg/dl(SD)) | Diferencias (95% IC) |
|-----------------|----------|-------------------------|--------------------------|------------------------------|
| COX5A | TT | 967 | 76.71 (0.34) | 0.00 |
| rs1133322 | CT-CC | 2516 | 75.25 (0.22) | -1.42 (-2.17 - -0.68) |
| HWE:0.97 | p-valor | | | 0.0002 |

Tabla 63. Asociación hallada con la PAD en la población de Di@bet.es.

Con respecto a los SNPs que se han asociado con PAS, el genotipo CC del rs1022580 y el GG del rs2746462 se relacionaron con valores más altos de este parámetro, aunque los individuos con dichos genotipos fueron muy escasos. Este último SNP es uno de los identificados en la fase 2. Los resultados en detalle se muestran en la tabla 64.

| dbSNP | Genotipo | Número de pacientes (N) | PAS (Media en mg/dl(SD)) | Diferencias (95% IC) |
|---------------|----------|-------------------------|--------------------------|------------------------------|
| rs1022580 | AA-CA | 3459 | 127.33 (0.31) | 0.00 |
| HWE: 1 | CC | 3 | 151 (21.2) | 28.74 (12.14 - 45.34) |
| | p-valor | | | 0.0007 |
| rs2746462 | TT-GT | 3389 | 127.36 (0.32) | 0.00 |
| HWE:1 | GG | 3 | 151 (21.2) | 28.74 (12.07 - 45.41) |
| | p-valor | | | 0.0007 |

Tabla 64. Asociaciones halladas con la PAS en la población de Di@bet.es.

5.4.6.1 Resultados de asociaciones de SNPs de CMR y HTA en subpoblaciones

5.4.6.1.1 Resultados de las asociaciones de CMR Y HTA según género

En la **población femenina**, no se ha obtenido ninguna relación con HTA ni con PAD. Sin embargo, se han encontrado asociaciones con PAS (tabla 65) y los mismos polimorfismos que se habían asociado en la población total. Al obtenerse asociaciones más significativas en mujeres, se podría decir que esta asociación está afectada por el género, pero siguen siendo muy escasos los pacientes con los genotipos asociados, a pesar de que el efecto es muy grande.

En la **población masculina**, no se ha hallado asociación con HTA ni con PAS, pero sí con el rs1133322 y PAD (tabla 66). Al verse una asociación más significativa en hombres que en la población general, se podría decir que el género influiría de forma muy importante en el efecto de este SNP sobre la PAD.

| dbSNP | Genotipo | Número de pacientes (N) | PAS (Media en mg/dl(SD)) | Diferencias (95% IC) |
|----------------------------|----------|-------------------------|-----------------------------|------------------------------|
| rs1022580 HWE: 1 | AA-CA | 2007 | 122.4 (0.4) | 0.00 |
| | CC | 2 | 164 (29) | 48.90 (29.06 - 68.73) |
| | p-valor | | | <0.0001 |
| rs2746462 HWE:1 | TT-GT | 1967 | 122.44 (0.41) | 0.00 |
| | GG | 2 | 164 (29) | 48.92 (29.01 - 68.82) |
| | p-valor | | | <0.0001 |

Tabla 65. Asociaciones halladas con PAS en la población femenina de Di@bet.es.

| dbSNP | Genotipo | Número de pacientes (N) | PAD (Media en mg/dl(SD)) | Diferencias (95% IC) |
|------------------------------|----------|-------------------------|-----------------------------|------------------------------|
| rs1133322 HWE:0.47 | TT | 432 | 79.61 (0.5) | 0.00 |
| | CT-CC | 1038 | 77.45 (0.3) | -2.23 (-3.28 - -1.17) |
| | p-valor | | | <0.0001 |

Tabla 66. Asociación obtenida con la PAD en la población masculina.

5.4.6.1.2 Resultados de las asociaciones de CMR Y HTA en individuos obesos

En la población obesa, se han obtenido asociaciones con PAS (tabla 67) y con PAD (tabla 68), pero no con HTA. El polimorfismo asociado con PAS no se ha asociado con otra población estudiada, por lo que podría decirse que para ver este comportamiento del SNP frente a PAS, la obesidad es un factor relevante. Por el contrario, la variación relacionada con PAD en obesos ya se había asociado de manera más significativa en la población total y en hombres, por lo que la obesidad no influiría decisivamente para producir esta asociación, aunque podría estar afectando de manera negativa, es decir, el efecto protector del alelo C no se podría manifestar en presencia de obesidad.

| dbSNP | Genotipo | Número de pacientes (N) | PAS (Media en mg/dl(SD)) | Diferencias (95% IC) |
|---------------------------|----------|-------------------------|-----------------------------|-----------------------|
| rs1136224 HWE:1 | AA | 566 | 135.29 (0.75) | 0.00 |
| | AG | 206 | 132.84 (1.19) | -3.03 (-5.46 - -0.60) |
| | GG | 18 | 127.1 (4.44) | -6.10 (-13.28 - 1.07) |
| | p-valor | | | 0.0045 |

Tabla 67. Asociación obtenida con la PAS en la población obesa de Di@bet.es.

| dbSNP | Genotipo | Número de pacientes (N) | PAD (Media en mg/dl(SD)) | Diferencias (95% IC) |
|------------------------------|----------|-------------------------|-----------------------------|-----------------------|
| rs1133322 HWE:0.47 | TT | 224 | 80.89 (0.7) | 0.00 |
| | CT-CC | 629 | 79.3 (0.39) | -1.69 (-3.18 - -0.21) |
| | p-valor | | | 0.026 |

Tabla 68. Asociación obtenida con la PAD en la población obesa de Di@bet.es.

5.4.6.1.3 Resultados de las asociaciones de CMR Y HTA en individuos no obesos

Al analizar la población no obesa, no se han hallado asociaciones con HTA. Sin embargo, se han asociado con PAS dos polimorfismos (tabla 69), que han coincidido con los relacionados en población general y femenina. Las asociaciones menos significativas las han mostrado en no obesos, por lo que podríamos inferir que la presencia de obesidad estaría contribuyendo en la aparición de esta asociación. Por otro lado, el rs1133322 se ha asociado con PAD (tabla 70) en todas las poblaciones analizadas en la fase 4, excepto la femenina. Esta asociación es más significativa en no obesos sólo al compararla con los obesos, hecho que podría significar que la obesidad estaría afectando el efecto del polimorfismo sobre PAD. Se han incluido los datos de la asociación del rs11205591 con PAD en la tabla 70, porque eran cercanos a la significancia, y en la fase 1 de este trabajo había asociación. Así, este polimorfismo se va a tener en cuenta en el análisis de asociación de la población total estudiada en este trabajo.

| dbSNP | Genotipo | Número de pacientes (N) | PAS (Media en mg/dl(SD)) | Diferencias (95% IC) |
|---------------------|----------|-------------------------|-----------------------------|------------------------------|
| rs1022580 HWE: 1 | AA-CA | 2608 | 124.92 (0.35) | 0.00 |
| | CC | 2 | 159 (34) | 31.34 (11.30 - 51.39) |
| | p-valor | | | 0.0022 |
| rs2746462 HWE:1 | AA-CA | 2556 | 124.96 (0.36) | 0.00 |
| | CC | 2 | 159 (34) | 31.30 (11.20 - 51.40) |
| | p-valor | | | 0.0023 |

Tabla 69. Asociaciones halladas con PAS en la población no obesa de [Di@bet.es](#).

| dbSNP | Genotipo | Número de pacientes (N) | PAD (Media en mg/dl(SD)) | Diferencias (95% IC) |
|--------------------------|----------|-------------------------|-----------------------------|------------------------------|
| rs1133322 HWE:0.94 | TT | 743 | 75.45 (0.37) | 0.00 |
| | CT-CC | 1887 | 73.9 (0.26) | -1.32 (-2.18 - -0.46) |
| | p-valor | | | 0.0026 |
| rs11205591 HWE:0.0065 | CC-CG | 2150 | 74.37 (0.24) | 0.00 |
| | GG | 129 | 75.54 (0.86) | 1.57 (-0.27 - 3.40) |
| | | | | 0.094 |

Tabla 70. Asociaciones halladas con PAS en la población no obesa de [Di@bet.es](#).

5.4.7 Validación de las asociaciones replicadas en las poblaciones

Después de encontrar en la población [Di@bet.es](#), las asociaciones de la fase preliminar que se replicaron, o tuvieron un comportamiento similar, con valores p cercanos al nivel de significación, éstas se estudiaron en la población obtenida a partir de la suma de los 4 estudios incluidos en este trabajo.

5.4.7.1 Asociaciones replicadas en las poblaciones generales totales

Al comparar las asociaciones halladas en la fase 1 y 3 en la población total, con los hallazgos en la muestra de [Di@bet.es](#), no se ha replicado ninguna asociación con las variables analizadas.

5.4.7.2 Comparación de las asociaciones con OB obtenidas en la fase 1 y 4

Asociaciones replicadas con OB en las poblaciones según género

En el caso de las poblaciones masculinas y femeninas, se ha visto que el alelo C del rs3770989 ha mostrado el mismo comportamiento con respecto a su asociación con OB, en hombres y en mujeres de las poblaciones de la fase 1 y 4. Sin embargo, en la

población total del estudio preliminar se había encontrado una asociación significativa en mujeres, mientras que en la población Di@bet.es estaba presente en hombres. Para estudiar más a fondo su comportamiento, se analizó este SNP en la población total de este trabajo, hallándose asociaciones significativas para ambas poblaciones, lo que confirma el hallazgo inicial. En la tabla 71 puede apreciarse que el efecto en hombres es protectoro, y de manera opuesta, en las mujeres presenta mayor riesgo de DMT2, para cada población independiente estudiada, y en su conjunto.

| dbSNP | Genotipos | No Obesidad (n(%)) | | Obesidad (n(%)) | | OR(95% CI) | |
|----------------------------|-----------|--------------------|-------------|-----------------|-------------|------------------------|------------------------|
| | | H | M | H | M | H | M |
| ASOCIACION FASE 1 | | | | | | | |
| NDUFS1 | TT | 1069 (92.2) | 1202 (93.6) | 364 (93.8) | 493 (89.3) | 1 | 1 |
| rs3770989 | CT-CC | 91 (7.8) | 82 (6.4) | 24 (6.2) | 59 (10.7) | 0.79(0.50-1.26) | 1.89(1.30-2.73) |
| HWE:1 | p-valor | | | | | 0.32 | 0.0009 |
| ASOCIACION FASE 4 | | | | | | | |
| NDUFS1 | TT | 930 (91.8) | 1338 (93.4) | 450 (94.7) | 568 (93.6) | 1 | 1.00 |
| rs3770989 | CT | 83 (8.2) | 94 (6.6) | 25 (5.3) | 39 (6.4) | 0.62(0.39-0.98) | 0.98(0.65-1.46) |
| HWE:0.5 | p-valor | | | | | 0.035 | 0.91 |
| ASOCIACION FASE 1+4 | | | | | | | |
| NDUFS1 | TT | 2025 (91.8) | 2601 (93.6) | 768 (94.7) | 1018 (91.6) | 1 | 1 |
| rs3770989 | CT-CC | 180 (8.2) | 179 (6.4) | 43 (5.3) | 93 (8.4) | 0.64(0.45-0.90) | 1.37(1.04-1.80) |
| HWE:0.2 | p-valor | | | | | 0.0072 | 0.026 |

Tabla 71. Asociaciones del SNP rs3770989 con OB en el estudio preliminar, en la réplica, y en el análisis conjunto de toda la población de este trabajo. H: hombres, M: Mujeres.

5.4.7.3 Comparación de las asociaciones con DMT2 obtenidas en la fase 1 y 4

Asociaciones replicadas con DMT2 en las subpoblaciones según género

En los subgrupos obtenidos al dividir las poblaciones totales según el sexo, se vio que el alelo G del rs9360898 presentaba la misma tendencia en DMT2 en la población masculina y femenina de la fase 1 y 4, a pesar de que se ha asociado significativamente con mayor riesgo en las mujeres de la fase 1 y con menor riesgo de OB en hombres de Di@bet.es. Para estudiar más a fondo este hallazgo, se analizó este polimorfismo en la población conjunta de las 2 fases. Así, se ha obtenido asociación significativa para las poblaciones femenina y masculina, en la población conjunta, corroborando el efecto opuesto de este polimorfismo sobre el riesgo de DMT2 en hombres y mujeres. En la tabla 72 se indican los resultados de estas asociaciones.

| Gen SNP | Genotipo | NO DMT2 (N%) | | DMT2 (N%) | | OR (95% CI) | |
|----------------------------|----------|--------------|-------------|------------|------------|-----------------------|-----------------------|
| | | H | M | H | M | H | M |
| ASOCIACION FASE 1 | | | | | | | |
| COX7A2 rs9360898 | TT | 770 (61.6) | 975 (62.1) | 123 (67.2) | 88 (51.5) | 1 | 1 |
| | GT-GG | 479 (38.4) | 594 (37.9) | 60 (32.8) | 83 (48.5) | 0.76/0.54-1.07 | 1.57/1.12-2.21 |
| | p-valor | | | | | 0.11 | 0.0091 |
| ASOCIACION FASE 4 | | | | | | | |
| COX7A2 rs9360898 | TT | 983 (61.4) | 1383 (59.7) | 230 (66.7) | 161 (57.9) | 1 | 1 |
| | GT-GG | 619 (38.6) | 932 (40.3) | 115 (33.3) | 117 (42.1) | 0.70/0.54-0.92 | 1.12/0.85-1.47 |
| | p-valor | | | | | 0.0087 | 0.44 |
| ASOCIACION FASE 1+4 | | | | | | | |
| COX7A2 rs9360898 | TT | 1764 (61.6) | 2385 (60.8) | 355 (66.2) | 256 (55.6) | 1 | 1 |
| | GT-GG | 1101 (38.4) | 1538 (39.2) | 181 (33.8) | 204 (44.4) | 0.75/0.61-0.92 | 1.25/1.02-1.55 |
| | p-valor | | | | | 0.0059 | 0.036 |

Tabla 72. Asociaciones del SNP rs9360898 con DMT2 en el estudio preliminar, en la réplica, y en el análisis conjunto de toda la población de este trabajo. H: hombres, M: Mujeres.

Con respecto a las asociaciones a replicar en glucemia, se ha encontrado una asociación significativa de la misma magnitud en las mujeres de ambas poblaciones generales analizadas, del alelo T del rs10504961 con valores altos de glucemia. Se ha validado su efecto en el análisis de la población femenina de la muestra total conjunta. En la tabla 73 se muestran los valores medios de glucemias y las diferencias entre los distintos genotipos, para los hombres y mujeres de las poblaciones de la fase 1 y 4, y para la población obtenida de su unión.

| Gen SNP | Geno- tipos | N | | Glucemia (Media en mg/dl(SD)) | | Diferencias (95%IC) | |
|---------------------------------|----------------|------|------|----------------------------------|--------------|----------------------------|---------------------------|
| | | H | M | H | M | H | M |
| ASOCIACION FASE 1 | | | | | | | |
| UQCRB rs10504961 HWE:0.43 | CC | 395 | 477 | 96.2 (1.12) | 90.47(0.66) | 0.00 | 0.00 |
| | CT-TT | 977 | 1163 | 94.0 (0.5) | 92.02(0.52) | -2.43 (-4.5 - -0.3) | 1.79 (0.09 - 3.49) |
| | p-valor | | | | | 0.026 | 0.039 |
| ASOCIACION FASE 4 | | | | | | | |
| UQCRB rs10504961 HWE:0.28 | CC | 470 | 708 | 96.84 (1.05) | 90.56 (0.58) | 0.00 | 0.00 |
| | CT-TT | 1166 | 1554 | 98.96 (0.7) | 91.76 (0.41) | 1.50 (-0.94 - 3.9) | 1.34 (0.06 - 2.63) |
| | p-valor | | | | | 0.23 | 0.041 |
| ASOCIACION FASE 1+4 | | | | | | | |
| UQCRB rs10504961 HWE:0.66 | CC | 724 | 1749 | 97.15 (0.88) | 93.74 (0.47) | 0.00 | 0.00 |
| | CT-TT | 1851 | 4238 | 97.68 (0.52) | 94.88 (0.3) | 0.14 (-1.74 - 2.0) | 1.14 (0.11 - 2.18) |
| | p-valor | | | | | 0.88 | 0.03 |

Tabla 73. Asociaciones del SNP rs3770989 con OB en el estudio preliminar, en la réplica, y en el análisis conjunto de toda la población de este trabajo. H: hombres, M: Mujeres.

Asociaciones replicadas con DMT2 en las subpoblaciones obesa y no obesa

De las asociaciones con DMT2 halladas en la **población obesa** en la fase 1, se ha encontrado en el Estudio Di@bet.es la misma tendencia para el SNP rs10742816, el alelo C se ha relacionado con mayor riesgo de esta enfermedad. Los análisis estadísticos han indicado una relación muy cercana a ser significativa, lo que apoya el efecto de ese polimorfismo sobre la DMT2. Al incrementar el tamaño muestral en el análisis de toda la población obesa de este trabajo, hemos encontrado una asociación significativa, confirmando el efecto de este polimorfismo sobre la DMT2. Los resultados para todas las poblaciones se muestran en la tabla 74.

| Gen/SNP | Genotipo | No DMT2 (n/%) | DMT2 (n/%) | OR (95 % CI) |
|----------------------------|----------|---------------|------------|---------------------------|
| ASOCIACION FASE 1 | | | | |
| NDUFS3 | TT-CT | 694 (99.9) | 181 (98.4) | 1.00 |
| rs10742816 | CC | 1 (0.1) | 3 (1.6) | 19.90(2.02-196.01) |
| HWE: 0.05 | p-valor | | | 0,0052 |
| ASOCIACION FASE 4 | | | | |
| NDUFS3 | TT-CT | 999 (98.9) | 307 (98.1) | 1.00 |
| rs10742816 | CC | 11 (1.1) | 6 (1.9) | 2.72 (0.94-7.84) |
| HWE: 0.75 | p-valor | | | 0,079 |
| ASOCIACION FASE 1+4 | | | | |
| NDUFS3 | TT-CT | 1628 (99.3) | 497 (98.4) | 1.00 |
| rs10742816 | CC | 12 (0.7) | 8 (1.6) | 3.26 (1.27-8.38) |
| HWE: 0.3 | p-valor | | | 0.02 |

Tabla 74. Asociaciones del SNP rs10742816 con DMT2 en el estudio preliminar, en la réplica, y en el análisis conjunto de toda la población obesa de este trabajo.

En la **población no obesa**, ha ocurrido algo similar a lo comentado anteriormente para los pacientes obesos. Las variaciones rs1136224 y rs2410718 se han asociado significativamente en la fase 1 con DMT2, mientras que en el estudio Di@bet.es han obtenido valores p muy cercanos al punto de corte a partir del cual se considera significativa una asociación. Una vez más, se han validado las asociaciones al obtener un resultado concordante en la población total no obesa (tabla 75).

5.4.7.4 Comparación de las asociaciones con HTA obtenidas en la fase 1 y 4

No se han encontrado asociaciones similares con HTA en las poblaciones femeninas y masculinas de las poblaciones analizadas en la fase 1 y en la fase 4.

| Gen/SNP | Genotipo | No DMT2 (n/%) | DMT2 (n/%) | OR (95 % CI) |
|----------------------------|----------|---------------|------------|--------------------------|
| ASOCIACION FASE 1 | | | | |
| NDUFS2 | AA | 1607 (76.8) | 126 (70.8) | 1 |
| rs1136224 | AG | 452 (21.6) | 46 (25.8) | 1.43 (0.99-2.06) |
| HWE: 0.62 | GG | 34 (1.6) | 6 (3.4) | 2.62 (1.03-6.68) |
| | p-valor | | | 0.013 |
| COX7C | GG | 1956 (92.5) | 153 (86.4) | 1 |
| rs2410718 | AG | 154 (7.3) | 22 (12.4) | 1.68 (1.02-2.76) |
| HWE: 0.57 | AA | 4 (0.2) | 2 (1.1) | 1.74 (1.12-2.71) |
| | p-valor | | | 0.019 |
| ASOCIACION FASE 4 | | | | |
| NDUFS2 | AA | 1889 (72) | 192 (71.9) | 1 |
| rs1136224 | AG | 668 (25.5) | 72 (27) | 1.25 (0.91-1.70) |
| HWE: 0.66 | GG | 65 (2.5) | 3 (1.1) | 0.40 (0.12-1.41) |
| | p-valor | | | 0.093 |
| COX7C | GG | 2664 (91.3) | 273 (90.4) | 1 |
| rs2410718 | AG | 250 (8.6) | 27 (8.9) | 1.33 (0.84-2.08) |
| HWE:0.57 | AA | 5 (0.2) | 2 (0.7) | 1.45 (0.95-2.19) |
| | p-valor | | | 0.092 |
| ASOCIACION FASE 1+4 | | | | |
| NDUFS2 | AA | 3621 (74) | 326 (71.3) | 1 |
| rs1136224 | AG | 1164 (23.8) | 118 (25.8) | 1.26 (1.00-1.60) |
| HWE: 0.29 | GG | 105 (2.1) | 13 (2.8) | 1.44 (0.76-2.71) |
| | p-valor | | | 0.034 |
| COX7C | GG | 4782 (91.8) | 438 (89.4) | 1 |
| rs2410718 | AG | 418 (8) | 48 (9.8) | 1.35 (0.96-1.89) |
| HWE: 0.57 | AA | 11 (0.2) | 4 (0.8) | 3.91 (1.05-14.51) |
| | p-valor | | | 0.021 |

Tabla 75. Asociaciones de los SNPs rs2410718 y rs1136224 con DMT2 en el estudio preliminar, en la réplica, y en el análisis conjunto de toda la población de este trabajo.

Asociaciones replicadas con HTA en las subpoblaciones obesa y no obesa

En la réplica en el estudio Di@bet.es, de los hallazgos obtenidos en el análisis de asociación con HTA de la fase 1, se ha encontrado una asociación con el parámetro PAD y el rs11205591, en la **población no obesa**, que ha mostrado la misma tendencia, acercándose su p-valor al punto de corte para considerarla como significativa. Después de realizar el análisis de asociación en la población total no obesa de este trabajo, se obtuvo la misma relación entre el SNP y la PAD, en este caso estadísticamente significativa. Los resultados para cada población se indican en la tabla 76.

| Gen dbSNP | Genotipo | Número de pacientes (N) | PAD (Media en mg/dl(SD)) | Diferencias (95% IC) |
|----------------------------|----------|-------------------------|--------------------------|---------------------------|
| ASOCIACION FASE 1 | | | | |
| NDUFS5 | CC-CG | 1898 | 76.31 (0.24) | 0.00 |
| rs11205591 | GG | 157 | 77.78 (0.79) | 1.62 (0.06 - 3.19) |
| | p-valor | | | 0.042 |
| ASOCIACION FASE 4 | | | | |
| NDUFS5 | CC-CG | | 74.37 (0.24) | 0.00 |
| rs11205591 | GG | 2150 | 75.54 (0.86) | 1.57 (-0.27 - 3.40) |
| | p-valor | 129 | | 0.094 |
| ASOCIACION FASE 1+4 | | | | |
| NDUFS5 | CC-CG | 3528 | 75.89 (0.18) | 0.00 |
| rs11205591 | GG | 243 | 77.06 (0.62) | 1.42 (0.11 - 2.73) |
| | p-valor | | | 0.034 |

Tabla 76. Asociaciones del SNP rs11205591 con PAD en el estudio preliminar, en la réplica, y en el análisis conjunto de toda la población no obesa de este trabajo.

5.4.7.5 Comparación de las asociaciones con OB, DMT2 Y HTA obtenidas en la fase 3 y 4

Del análisis estadístico realizado con las variantes identificadas en la fase 2, en la población total de la fase 1, se han obtenido las asociaciones del estudio de la fase 3. En la población Di@bet.es, no se han hallado resultados concordantes con los obtenidos en la fase 3.

6. Discusión

La DMT2 es una enfermedad compleja, heterogénea y multifactorial en la cual intervienen factores genéticos, ambientales y adquiridos. El conjunto de loci y variantes genéticas que intervienen en la enfermedad están relacionados principalmente con la disfunción de las células β pancreáticas y con alteraciones metabólicas como la resistencia a la insulina. El efecto de los diferentes genes implicados, junto con factores ambientales, como la obesidad, la hipertensión, modulan el umbral para que se desarrolle o no la enfermedad (Langenberg et al 2014). Diversos estudios muestran la importancia de la heredabilidad de la DMT2, estableciendo que posiblemente es superior al 70% (Medici et al y Poulsen et al 1999). Sin embargo, pese a todos los estudios realizados, el conjunto de loci y variantes genéticas identificadas explican únicamente entre el 10 y 15% de la heredabilidad de la enfermedad (Zeggini et al 2008, Billings et al 2010). Algo similar sucede en otras enfermedades complejas, como la obesidad (Pigeyre et al 2016) y la hipertensión (Zhao et al 2013).

Son varias las estrategias que se han seguido para conocer que genes están asociados a la DMT2 (Florez 2008 y Bonnefond et al 2010): estudio de genes candidatos, estudio de ligamiento de familias, GWAS, entre otras. El mayor avance se ha visto con la llegada de los estudios GWAS (<http://www.ebi.ac.uk/gwas>). Sin embargo, en ellos no se han estudiado específicamente variantes de los genes implicados en distintos sistemas biológicos, como la cadena mitocondrial respiratoria, a pesar de que en estos últimos años, las distintas plataformas disponibles en el mercado (Affimetrix, Illumina) han aumentado en sus arrays el número de marcadores. Por otro lado, no hay estudios de este tipo que incluyan población española.

El presente trabajo surge dentro de un proyecto coordinado por varios grupos de investigación españoles, con el objetivo principal de caracterizar la DMT2 en España, siendo el objetivo de nuestro grupo, la identificación de variantes genéticas presentes en la cadena mitocondrial respiratoria asociadas con el desarrollo de la DMT2, en una muestra representativa de la población española.

Características generales de las poblaciones del estudio

Los estudios poblacionales incluidos en este trabajo han mostrado heterogeneidad con respecto a sus características antropométricas, bioquímicas y clínicas. La población de Pizarra ha sido la que ha presentado mayores valores en el porcentaje de DMT2, OB, OB Central y HTA. En concordancia con esto, ha mostrado las medias más elevadas para el IMC, cintura, glucemia, PAS y PAD. Sin embargo, es la población con la menor edad media, al contrario de lo que se podría pensar, según el efecto que produce el aumento de la edad sobre todos estos parámetros. Además, es la que más proporción de mujeres ha incluido. En los estudios Horteiga y Segovia se han visto valores más similares para todos los parámetros evaluados, aunque aun así, se han obtenido diferencias significativas entre sus medias y porcentajes, excepto para HTA y cintura. A pesar de estos rasgos poblacionales, se han obtenido asociaciones y tendencias similares en las 3 poblaciones analizadas individualmente con muchos polimorfismos, que luego han confirmado su asociación al juntar los 3 estudios en una población conjunta.

En el estudio preliminar de los polimorfismos de la CMR, se han agrupado los individuos de los estudios Horteiga, Pizarra y Segovia, creando así una muestra más amplia de población española. De esta manera, se ha aumentado el tamaño muestral, lo que ha permitido estudiar variaciones menos frecuentes y minimizar la posibilidad de obtener asociaciones al azar, aumentando el poder estadístico.

La población utilizada para la réplica de las asociaciones halladas inicialmente, ha sido la del primer estudio realizado a nivel nacional sobre DMT2 y alteraciones del metabolismo glucídico en España, llamado Di@bet.es. En ella, han participado individuos de todas las comunidades autónomas españolas, excepto de La Rioja. Las comunidades más representadas son en primer lugar Andalucía, seguidos de Cataluña y la Comunidad de Madrid. De la provincia de Málaga, de donde se recolectaron las muestras del estudio Pizarra, sólo se incluyeron 230, mientras que de las provincias de Valladolid y Segovia, a donde pertenecen los participantes de los estudios Horteiga y Segovia respectivamente, no se incluyó ninguno. La

comunidad autónoma a la que pertenecen estas dos últimas provincias, Castilla y León, sólo está representada por 357 personas. La distribución de muestras en este estudio se ha basado en el número de individuos de cada comunidad, para que sea representativo de toda España.

En el caso de la comparación entre la población total del análisis preliminar, y el estudio utilizado para la réplica de sus resultados, las diferencias son significativas para todos los parámetros que se han tenido en cuenta a la hora de definir a las poblaciones, a excepción de la edad y cintura. Los valores más elevados los ha mostrado el estudio Di@bet.es, excepto en el caso de la PAD y la HTA. Estos hallazgos proceden de las diferencias ya descritas entre las poblaciones analizadas, y de otras características, como el hecho de que las muestras mayoritarias del estudio Di@bet.es pertenecen a provincias con más población y mayor migración, mientras que las poblaciones Pizarra, Valladolid y Segovia no comparten ese rasgo.

Asociaciones entre la CMR y riesgo de DMT2

En este estudio hemos hallado asociaciones entre varios genes de la CMR y las enfermedades DMT2, OB y HTA. Se han estudiado de forma preliminar, varias poblaciones de distintas regiones españolas de forma independiente y también en su conjunto. Además, se han definido subgrupos de las poblaciones generales según el sexo y la presencia o no de obesidad, para analizar el comportamiento de los SNPs de la CMR en ellos. Por otro lado, se han identificado variaciones en los genes asociados en la primera fase, que pudieran relacionarse con estas enfermedades. A continuación, se han analizado los mismos polimorfismos en otra población general española. Finalmente, las variantes replicadas, se han analizado en la población total de este trabajo.

Una forma más estricta de evaluar los resultados de asociación obtenidos en comparaciones múltiples, es a través del uso de la corrección por Bonferroni. Esta supone un cambio del nivel de significación α para minimizar el error de encontrar una asociación por azar. Sin embargo, según el diseño de nuestro trabajo, no es

necesario que las asociaciones sean significativas según esta modificación, debido a que las asociaciones halladas, se van a validar a través de la réplica en otra población. Esta discusión se centrará en las asociaciones más relevantes en las distintas etapas del trabajo, para las distintas variables y poblaciones analizadas.

Asociaciones halladas en la primera fase (Fase1)

En el estudio preliminar de este trabajo, las asociaciones halladas en las poblaciones individuales, han sido significativas sólo en una de las tres poblaciones. Sin embargo, las poblaciones restantes cumplían la misma tendencia. En general, las asociaciones más significativas se hallaron en el Estudio Pizarra, lo que puede explicarse por el hecho de que esa población tiene los porcentajes más altos de obesidad, diabetes e hipertensión, y de los parámetros que definen dichas patologías. De todas formas, al seleccionar esos polimorfismos para ser estudiados en la población conjunta, se han podido obtener asociaciones más significativas con DMT2 en la muestra total. Si bien, la significación y las diferencias entre los genotipos no han sido muy relevantes, se han reforzado las relaciones obtenidas entre esos polimorfismos y la DMT2.

En el estudio de la población conjunta, el alelo G del polimorfismo rs10891319 (SDHD), se ha asociado con valores altos de IMC y cintura en la población general. Además, existe una asociación de ese alelo con mayor riesgo de OB, lo que indicaría que la modificación del IMC estaría afectando al riesgo de OB. Esta variación se encuentra en el extremo 3' del gen, y podría estar mediando su efecto a través de la regulación de la expresión del gen al que pertenece, interfiriendo por ejemplo en la acción de algún factor específico de transcripción, o en las interacciones de largo alcance en la cromatina (Li et al 2016), o estar en desequilibrio de ligamiento con otro polimorfismo responsable de este efecto. No se ha encontrado hasta el momento ninguna asociación con este SNP y la obesidad. Sólo se ha obtenido una asociación de este gen con el espesor de grasa de espalda en un estudio de genes candidatos relacionados con obesidad en cerdos (Fontanesi et al 2012). Otro polimorfismo con una asociación ha sido el rs3770989 (NDUFS1), cuyo alelo C ha mostrado también valores altos de IMC. Sin embargo, es en el caso de la población femenina, donde la

asociación se vuelve mucho más significativa, y además, se relaciona con el riesgo de OB. Esto podría indicar que las características distintivas de esta subpoblación, como por ejemplo el perfil hormonal, podrían estar favoreciendo el efecto de este alelo sobre la obesidad. El mecanismo por el cual este SNP podría afectar la acción del gen, sería a través de la regulación de la expresión, debido a que se encuentra en la región 3' UTR y podría así modificar el sitio de unión de miARNs, o estar en desequilibrio de ligamiento con el SNP causante del efecto. El único trabajo que relaciona a NDUFS1 con obesidad es el realizado por Sparks et al en 2005, en el que han comprobado en ratones, que la expresión del gen NDUFS1 estaba disminuida en músculo esquelético, al exponerlos a una dieta alta en grasa.

El gen SDHC, se ha asociado con todos los parámetros relacionados con obesidad estudiados, lo que sugiere un efecto importante sobre esta enfermedad. El alelo G del polimorfismo rs4600063, ha mostrado valores bajos de IMC y cintura, a la vez que se ha asociado con menor riesgo de obesidad y obesidad central. La presencia de este alelo parece ser protectora para el desarrollo de obesidad. Al analizar las poblaciones masculina y femenina, sólo se ha asociado con obesidad central en mujeres, mejorando su valor p. Por su localización en la región 3' UTR, podría mediar su efecto de manera similar al polimorfismo descrito anteriormente. Ningún estudio ha relacionado este SNP o gen con obesidad, aunque sí se ha relacionado con DMT2 a SDHC (Wu et al 2017).

Con respecto al análisis de asociación con la DMT2, el genotipo GG del rs11205591 (NDUFS5) ha mostrado menor riesgo de DMT2 en la población total, con lo cual la herencia recesiva del alelo G podría estar teniendo un efecto protector contra la enfermedad. Esta variación se encuentra en el extremo 3' del gen, y podría estar regulando la expresión de éste. No se ha encontrado hasta el momento ninguna asociación con este gen y la diabetes. En el caso del alelo A del rs17624157 (UQCRCQ), éste se ha relacionado con mayor riesgo de DMT2 sólo en hombres, lo que indicaría que el género interactúa con el alelo para favorecer su efecto. La presencia de obesidad parece ser otro factor que influye en la asociación de este alelo con el desarrollo de la DMT2, debido a que aunque se ha asociado con valores

altos de glucemia en la población general, se ha asociado más significativamente en obesos. En este caso, la obesidad parece estar afectando el comportamiento de esta variación con respecto a la glucemia, y no de forma directa al riesgo de DMT2. Este SNP se encuentra en la región 3' UTR, por lo que podría regular la expresión del gen al modificar sitios de unión de miARNs. No se ha relacionado este gen en ningún trabajo previo con la DMT2. Estos polimorfismos podrían estar en desequilibrio de ligamiento con otros, que sean los responsables del efecto encontrado.

En el estudio de la HTA, se ha visto la asociación del alelo C del rs1053517 (NDUFS1) con valores altos de PAD en población general y en la población masculina. Sin embargo, la asociación más fuerte la ha presentado en hombres, lo que sugiere que el género influye en la aparición del efecto descrito. Por su localización en la región 3' del gen, este polimorfismo podría regular de alguna manera la expresión del gen, o podría ser que otro SNP en desequilibrio de ligamiento, sea el causante del efecto. Esta variación se ha relacionado con ganancia de peso inducida por antipsicóticos en pacientes con esquizofrenia (Gonçalves et al 2014), pero no con hipertensión. Un estudio reciente ha concluido que la expresión del gen NDUFS1 está disminuida en músculo esquelético de diabéticos (Wu et al 2017). Por otro lado, la dieta rica en grasas también provocaba este efecto en músculo esquelético de ratones (Sparks et al en 2005). Sin embargo, no se ha visto asociación de este gen con la hipertensión arterial.

Identificación y estudio de variantes genéticas en genes asociados (Fase 2 y 3)

Al analizar el gen NDUFS5, el rs11205591 que se ha utilizado para dividir controles y casos a estudiar según sus genotipos, está en desequilibrio de ligamiento con los SNPs con frecuencias alélicas diferentes identificados en la fase 2, excepto con el SNP rs6981. Como era de esperar considerando esto, todos han mostrado asociación con OB y DMT2. En el caso del gen NDUFS1, se seleccionaron los individuos según los genotipos del rs3770989. Este SNP está en desequilibrio de ligamiento con el rs77655991, pero no con los otros 3 SNPs estudiados de este gen, que sí lo hacen entre ellos. Por esta razón, se podría explicar el comportamiento distinto de los SNPs rs4147718, rs1801318 y rs4147729, que no han mostrado asociación con

OB y sí con DMT2, mientras que el rs77655991 se ha asociado con OB pero no con DMT2, de manera similar al rs3770989 en la fase 1. Los polimorfismos del gen SDHB estaban en desequilibrio de ligamiento y ambos se han asociado con OB.

El gen NDUFS1 ha mostrado una asociación fuertemente significativa con HTA. El alelo C del rs1801318 se ha relacionado con un mayor riesgo de HTA en la población obesa. Este polimorfismo se ha asociado con la enfermedad solamente en presencia de obesidad, por lo que este desorden podría ser un factor importante para que se produzca el efecto de este alelo, y favorezca la aparición de la HTA. Este SNP es una variante sinónima, variaciones que en la actualidad se sabe que pueden estar afectando la expresión, conformación o función de la proteína (Sauna et al 2011), aunque su efecto podría estar causado por otro polimorfismo que esté en desequilibrio de ligamiento. Esta variación se ha asociado previamente con ganancia de peso inducida por antipsicóticos en pacientes con esquizofrenia (Gonçalves et al 2014), pero no con hipertensión.

Estudios de replicación de las asociaciones (Fase4)

Asociaciones halladas en la población Di@bet.es

Al estudiar el comportamiento frente a la DMT2, de las variantes seleccionadas, el alelo C del rs1470625 (COX5B) se ha asociado con mayor riesgo de la enfermedad en la población total y en la femenina, siendo en esta última donde muestra una asociación más fuerte. Este polimorfismo se ubica en la región 3' del gen, por lo que podría afectar sitios de unión de factores de transcripción o de interacción cromatínica, reguladores de la traducción, unión de miARNs, o estar en desequilibrio de ligamiento con polimorfismos realmente asociados al efecto observado. Sólo se han relacionado hasta el momento 2 polimorfismos del gen COX5B con la DMT2 y otros desórdenes metabólicos de la glucosa, en el estudio de Snogdal et al del 2012, que no se han analizado en nuestro trabajo. En la población obesa, el genotipo GG del rs4806187 (COX6B1) ha mostrado valores elevados de glucemia. La asociación sólo se ha encontrado en obesos, por lo que la obesidad podría ser necesaria para que se produzca el efecto de este polimorfismo sobre la glucemia. Además, en no obesos este genotipo presenta glucemias más bajas

respecto al resto, sin llegar a mostrar una asociación. Su efecto podría medirse, de manera similar a las demás variantes descritas anteriormente, presentes en la región 3' del gen. El trabajo de Snogdal et al del 2012 también ha estudiado este gen, pero no ha hallado asociación con la DMT2.

Con respecto a la HTA, 3 variaciones se han relacionado muy significativamente con las presiones arteriales, una con PAD y el resto con PAS. El alelo C del rs1133322 (COX5A) se ha asociado con valores bajos de PAD en población total y en hombres. No obstante, la asociación más fuerte se produjo en la población masculina, lo que sugiere un efecto del género en la presentación de esta asociación. El gen al que pertenece, se ha asociado a disfunción del islote pancreático en ratones diabéticos y controles (Lu et al 2008). Los 2 polimorfismos asociados con PAS, rs1022580 y rs2746462, pertenecen al gen SDHB, y están en desequilibrio de ligamiento, lo que podría explicar su comportamiento similar. El genotipo CC del primer SNP, y el GG del segundo, muestran valores más altos de PAS en la población general y en mujeres, siendo un polimorfismo cuyo efecto podría depender de género. La base de datos del NCBI, da una significación clínica benigna para el rs1022580, basándose en la base de datos que recopila las variaciones genómicas relacionadas con enfermedades humanas (ClinVar), y no muestra relación con ninguna enfermedad específica. Al ser un polimorfismo intrónico, el mecanismo por el cual podría afectar al gen SDHB, sería la modificación de sitios de splicing, entre otros, o estar en desequilibrio de ligamiento con otro polimorfismo realmente relacionados con el efecto hallado. Asimismo, para el rs2746462, la significancia clínica en el ClinVar es también benigna, aunque se ha encontrado en individuos con los fenotipos de paraganglioma y sarcoma estromal gástrico, feocromocitoma y síndrome hereditario predisponente al cáncer. Este último polimorfismo se ubica dentro del exón 1, pero no produce un cambio de aminoácido. Este polimorfismo podría relacionarse con la regulación del splicing, eficiencia de la transcripción y traducción, o estar en desequilibrio de ligamiento con otros polimorfismos causantes del efecto observado. Se ha demostrado de manera fehaciente, que muchas mutaciones en el gen SDHB son causales de enfermedades como el feocromocitoma o paraganglioma. Uno de los síntomas principales de estas

patologías es la presión arterial alta, lo que concuerda con el hecho de que polimorfismos estudiados en este trabajo, pertenecientes a este gen, se hayan asociado con la presión arterial sistólica. Por otro lado, existe una relación comprobada entre el feocromocitoma y la diabetes mellitus, a través de la acción de las catecolaminas en la célula beta del páncreas, que produce una tolerancia alterada de la glucosa por inhibición de la liberación de insulina (Nestler y McClanahan 1992, Adlan et al 2010).

Asociaciones que se han replicado en este estudio

El último paso de este trabajo, fue validar las asociaciones identificadas, mediante un análisis de asociación en la población conjunta del estudio preliminar (fase1) y de la población de Di@bet.es, de los polimorfismos replicados.

Al comparar qué polimorfismos se habían asociado con las distintas variables analizadas, en la población conjunta de los estudios Hortega, Segovia y Pizarra, y en la población Di@bet.es, sólo un polimorfismo ha mostrado exactamente la misma asociación, con idéntica significación. Este es el rs10504961 del gen UQCRB, cuyo alelo T se ha asociado con valores elevados de glucemia en la población femenina. A pesar de que esta asociación no era muy significativa, en el análisis realizado en la población total de este trabajo, se ha confirmado esta asociación. Su efecto es moderado si consideramos las diferencias de glucemias de los distintos genotipos, lo que es característico de polimorfismos con MAFs altos. El SNP está ubicado en la región 3' UTR, por lo que podría modular la expresión del gen al que pertenece, modificando los sitios de unión de miARNs, u otro SNP en desequilibrio de ligamiento con él podría ser el responsable real de la asociación. Este gen no se ha relacionado anteriormente con la DMT2.

En el caso de las poblaciones obesa y no obesa, se han obtenido relaciones cercanas al nivel de significación en la población Di@bet.es, en asociaciones que eran significativas en la primera fase del trabajo. Un nuevo análisis en la población total incluida en este trabajo, ha confirmado todas estas asociaciones, aumentando el tamaño muestral se ha obtenido una asociación significativa.

Para los obesos, el genotipo CC del rs10742816 (NDUFS3) se asociaba con mayor riesgo de DMT2 en la fase 1, mientras que en la población utilizada en la réplica, a pesar de mostrar la misma tendencia, sólo mostraba un valor p muy cercano al valor de corte de significación. En el análisis de asociación de la población total, se ha corroborado esta relación, obteniéndose un efecto fuerte sobre el riesgo de la DMT2. Este SNP se ubica en la región 5', por lo que su presencia en sitios de unión de factores de transcripción podría modificar la expresión del gen. Existen 2 trabajos en los que se ha relacionado a este gen, uno con la diabetes mellitus tipo 1 (DMT1), y otro con la Diabetes mellitus gestacional (DMG). En el estudio realizado por Akude et al en 2011, se ha visto una menor producción de la proteína codificada por el gen NDUFS3 en ratas con DMT1, acompañada por una descenso en la actividad de los complejos mitocondriales. Además, Taurino et al en el 2012 han detectado una disminución del nivel de mRNA de ese gen, con una actividad catalítica alterada en ratas diabéticas. En embarazadas obesas con DMG se ha visto una reducción en el contenido de las subunidades del complejo mitocondrial I codificadas por NDUFS3 y NDUFV2, con una reducción en la actividad enzimática de este complejo (Boyle et al 2014).

De manera similar, en los no obesos, el alelo G del rs1136224 (NDUFS2) y el alelo A del rs2410718 (COX7C) se han asociado con mayor riesgo de DMT2 en la primera fase, pero sólo están cercanos al nivel de significación en la réplica. En la población obesa, el alelo G del rs1136224 tiene una tendencia opuesta respecto al riesgo de la DMT2, mientras que para el alelo A del rs2410718, su OR medio es cercano a 1. Las dos asociaciones se han confirmado en el análisis de validación. En ambos polimorfismos, pocos individuos no obesos con DMT2 presentaban el alelo asociado, y los efectos sobre la enfermedad son altos si evaluamos el OR, más aún para el rs2410718. El primer SNP está localizado en la región 3'UTR, por lo que podría afectar la unión de miARNs y así regular la expresión de NDUFS2. El rs2410718 se encuentra 3' del gen al que pertenece, por lo que su posible acción sería modificar sitios de unión de factores específicos de transcripción o de interacción cromatínica de largo alcance. Ambos polimorfismos podrían estar en desequilibrio de ligamiento con otros polimorfismos que estuvieran provocando esta

asociación. El rs1136224 se ha asociado con esclerosis múltiple en bibliografía (Ban et al 2008), pero no con DMT2. El gen COX7C presenta una reducción en los niveles de mRNA en pacientes obesos y obesos con DMT2 (Pihlajamäki et al, 2009).

Al analizar la HTA, se ha encontrado en la población no obesa, una asociación del genotipo GG de rs11205591 (NDUFS5) y mayores valores de PAD, que ha sido significativa en la población del estudio preliminar, pero sólo cercana al valor de corte en la población de Di@bet.es. Si miramos su comportamiento en la población obesa, no muestra ningún efecto sobre PAD. Al aumentar el tamaño muestral en el análisis de la población conjunta de este trabajo, se ha validado la relación, con un efecto modesto sobre los valores de PAD. Al encontrarse la relación con PAD en los no obesos, la obesidad podría estar influyendo de alguna manera en estas asociaciones. No se han hallado asociaciones de este gen con HTA en bibliografía hasta el momento.

En resumen, se han hallado 17 asociaciones significativas con genes de la CMR en las distintas poblaciones analizadas, para los distintos parámetros relacionados con la DMT2. Además, en la réplica de las asociaciones presentes en la primera fase, 5 polimorfismos han mostrado un comportamiento similar en la nueva población estudiada, y se han validado en el análisis de la población total de este trabajo. En población general española, se han obtenido asociaciones con OB y los genes SDHD y SDHC, y una asociación con DMT2 y NDUFS5 en el estudio preliminar. En la población femenina (tabla A) y en la masculina (tabla B), se han encontrado las siguientes asociaciones para las distintas etapas de este trabajo:

| Variable | Fase 1 | Fase 3 | Fase 4 | Replicadas |
|----------|---------------------------------------------|--------|-------------------------------------------------------------------|---------------------------------------------|
| OB | NDUFS1 ⁺⁺⁺ | - | NDUFS1 ⁰ | NDUFS1 ⁺⁺ |
| DMT2 | UQCRB ⁽⁺⁺⁾ /COX7A2 ⁺⁺ | - | COX5B ⁽⁺⁺⁾ /UQCRB ⁽⁺⁺⁾ /COX7A2 ⁰ | UQCRB ⁽⁺⁺⁾ /COX7A2 ⁺⁺ |
| HTA | - | - | SDHB ⁽⁺⁺⁾ | - |

Tabla A. Genes asociados a DMT2 y parámetros relacionados estudiados en las distintas etapas del trabajo, en la población femenina. ⁰ indica que tiene la misma tendencia, ⁽⁺⁾ indica que el p-valor es cercano a nivel de significación, ⁽⁺⁺⁾ indica p-valor < 0.05, y ⁽⁺⁺⁺⁾ indica que la asociación hallada tiene un p-valor < 0.001.

| Variable | Fase 1 | Fase 3 | Fase 4 | Replicadas |
|----------|--------------------------------------------|--------|------------------------|----------------------|
| OB | NDUFS1 ⁰ | - | NDUFS1 ⁺⁺ | NDUFS1 ⁺⁺ |
| DMT2 | UQCRQ ⁽⁺⁺⁾ /COX7A2 ⁰ | - | COX7A2 ⁺⁺ | COX7A2 ⁺⁺ |
| HTA | NDUFS1 ⁺⁺⁺ | - | COX5A ⁽⁺⁺⁺⁾ | - |

Tabla B. Genes asociados a DMT2 y parámetros relacionados estudiados en las distintas etapas del trabajo, en la población masculina. ⁰ indica que tiene la misma tendencia, ⁽⁺⁾ indica que el p-valor es cercano a nivel de significación, ⁽⁺⁺⁾ indica p-valor < 0.05, y ⁽⁺⁺⁺⁾ indica que la asociación hallada tiene un p-valor < 0.001.

En los análisis realizados en los individuos obesos (tabla C) y no obesos (tabla D) se han obtenido las asociaciones que se indican a continuación:

| Variable | Fase 1 Obesos | Fase 3 Obesos | Fase 4 Obesos | Replicadas Obesos |
|----------|-----------------------------------------------|-----------------------|-----------------------------------------------|------------------------|
| DMT2 | UQCRQ ⁽⁺⁺⁺⁾ /NDUFS3 ⁽⁺⁾ | - | COX6B1 ⁽⁺⁺⁾ /NDUFS3 ⁽⁺⁾ | NDUFS3 ⁽⁺⁺⁾ |
| HTA | | NDUFS1 ⁺⁺⁺ | | - |

Tabla C. Genes asociados a DMT2 y parámetros relacionados estudiados en las distintas etapas del trabajo, en la población obesa. ⁰ indica que tiene la misma tendencia, ⁽⁺⁾ indica que el p-valor es cercano a nivel de significación, ⁽⁺⁺⁾ indica p-valor < 0.05, y ⁽⁺⁺⁺⁾ indica que la asociación hallada tiene un p-valor < 0.001

| Variable | Fase 1 No Obesos | Fase 3 No Obesos | Fase 4 N Obesos | Replicadas No Obesos |
|----------|-----------------------------------------------|------------------|---------------------------------------------|-----------------------------------------------|
| DMT2 | COX7C ⁽⁺⁺⁾ /NDUFS2 ⁽⁺⁺⁾ | - | COX7C ⁽⁺⁾ /NDUFS2 ⁽⁺⁾ | COX7C ⁽⁺⁺⁾ /NDUFS2 ⁽⁺⁺⁾ |
| HTA | NDUFS5 ⁽⁺⁺⁾ | | NDUFS5 ⁽⁺⁾ | NDUFS5 ⁽⁺⁺⁾ |

Tabla D. Genes asociados a DMT2 y parámetros relacionados estudiados en las distintas etapas del trabajo, en la población no obesa. ⁰ indica que tiene la misma tendencia, ⁽⁺⁾ indica que el p-valor es cercano a nivel de significación, ⁽⁺⁺⁾ indica p-valor < 0.05, y ⁽⁺⁺⁺⁾ indica que la asociación hallada tiene un p-valor < 0.001

El hecho de que ciertas asociaciones se hayan encontrado sólo en hombres o en mujeres, estaría sugiriendo la influencia del género sobre ellas. De manera similar, las asociaciones que sólo se han obtenido en presencia o en ausencia de obesidad, estarían afectadas por este factor. Con respecto a las variaciones que se han asociado, 7 se han encontrado en la región 3' (4 asociadas con DMT2, 2 con HTA y una con OB), 6 en el extremo 3'UTR (4 con DMT2 y 2 con OB), 2 han sido variantes sinónimas (asociadas con HTA en obesos), una intrónica (con HTA), y otra en el extremo 5' (con DMT2). El mecanismo por el cual los efectos hallados

podrían medirse, no está claro, debido a que no se han hallado variaciones funcionales evidentes que se asocien con el desarrollo de la DMT2. Los polimorfismos asociados podrían estar en desequilibrio de ligamiento con polimorfismos realmente responsables del efecto observado. Al realizar una búsqueda bibliográfica de los genes y polimorfismos asociados, no se han encontrado publicaciones que los hayan relacionado con el parámetro con el que se asoció en este trabajo.

Estudio de variantes de la CMR en las plataformas GWAS

Muchos son los trabajos que se han realizado con el objetivo de identificar qué genes están implicados en la DMT2 y desórdenes relacionados. No obstante, con la llegada de los GWAS, se han obtenido grandes avances en el conocimiento de variantes involucradas en este y otros desórdenes complejos. Hasta este momento, se han estudiado variaciones comunes, con frecuencias de más del 5%. Sin embargo, ya existen arrays que incluyen variantes con MAF mayor al 1%, lo que creen conducirá a una nueva generación de GWAS. A pesar de la incorporación de esta mejora, el tamaño muestral para detectar estos marcadores debe ser muy alto. A su vez, otra limitación de estos estudios, es la cobertura de cada gen a analizar, que aunque se ha mejorado mucho, el aumento del número de polimorfismos estudiados, estará limitado por la densidad de éstos en cada región del genoma, es decir, si se concentran en una determinada región, no todas esas variantes podrán ser evaluadas, según el proceso de diseño de la técnica de los arrays.

Los primeros arrays en el mercado (318K, CNV-370, 300), incluían unas 331 variantes presentes en genes de la CMR, y abarcaban 24 de los 29 genes estudiados en este trabajo, mientras que en los arrays de nueva generación, han ampliado el total de marcadores a 874 (familia de arrays Human Omni), y ahora incluyen los 29 genes. Hasta hace unos años, sólo 6 SNPs de los arrays GWAS coincidían con los de nuestro estudio, incluido el SNP rs4806187, que sí ha mostrado asociación con la DMT2 en la población obesa de nuestro trabajo. En los arrays de nueva generación, son 20 los polimorfismos que coinciden con los analizados en este trabajo. De ellos,

4 se han asociado, el rs1136224 con DMT2, OB y HTA, el rs4656993 con DMT2, el rs1022580 con PAS, y el rs4600063 con OB.

Estudios de asociación GWAS con la DMT2, OB e HTA

El National Human Genome Research Institute, junto con el European Bioinformatics Institute (NHGRI-EBI) ha creado un catálogo de las publicaciones GWAS (<http://www.ebi.ac.uk/gwas>), relacionadas con las enfermedades de los distintos sistemas y aparatos del cuerpo humano. Se ha realizado una revisión de todas las asociaciones obtenidas para la DMT2, enfocada en los genes involucrados en la cadena mitocondrial respiratoria. De 751 variantes totales asociadas, sólo 2 pertenecían a este sistema biológico. Una de las variantes era el rs13253073 del gen COX6C, en la que el alelo A se ha asociado a menor riesgo de DMT2 (Palmer et al 2014). El otro polimorfismo asociado fue el rs2055942 del gen COX7B2 (Al Safar HS et al 2013). A pesar de que estos genes han sido incluidos para su estudio en nuestro trabajo, no se han analizado estos polimorfismos, ni se han hallado asociaciones con esos genes. En el caso de la OB, se han obtenido 1905 variaciones totales asociadas. Sin embargo, sólo 6 genes de la CMR se han relacionado significativamente en estudios GWAS. Estos son el COX15 (rs2281636), NDUFB8P2 (rs1051424), NDUFA10 (rs4852140), NDUFB4 (rs9289146), NDUFA8 (rs3818638) y UQCRHP1 (rs2844479). Ninguno de estos genes se ha analizado en este trabajo. Con respecto a la HTA, a pesar de haber 211 marcadores asociados, ninguno pertenece a la CMR.

Los estudios GWAS realizados hasta el momento con DMT2, OB y HTA, no han llegado a utilizar los arrays llamados de nueva generación descritos anteriormente, por lo que las asociaciones halladas en este trabajo, correspondientes a los polimorfismos incluidos en ellos, no se pueden cotejar con las obtenidas hasta el momento en esta clase de estudios. La excepción es la asociación del rs4806187 con la DMT2 en la población obesa, debido a que ese polimorfismo estaba incluido en los arrays más utilizados en GWAS. Sin embargo, los estudios de marcadores de genoma completo, en su mayoría, analizan cohortes de casos y controles, o

poblaciones generales, pero no específicamente una subpoblación, como en el caso de nuestra asociación, que se presenta en la población obesa. Otro hecho importante, es que no se han incluido en este tipo de estudios, a poblaciones españolas, y no hay muchos trabajos que analicen el comportamiento de estos marcadores en población mediterránea.

Metodologías de genotipado empleadas

El sistema SNPlex se ha utilizado en la fase preliminar de nuestro trabajo. Esta es una técnica adecuada para una primera aproximación del estudio de variaciones en poblaciones, debido a que es de alto rendimiento, es decir, se pueden determinar 48 polimorfismos simultáneamente en cada individuo, reduciendo así el volumen de muestra a emplear, así como el tiempo de procesamiento necesarios para completar los experimentos de gran tamaño muestral. Su especificidad viene dada en el paso de la ligación con sondas, y la sensibilidad está definida por la realización de la PCR. En nuestro laboratorio se han realizado muchos trabajos mediante esta técnica (Mansego et al 2011, 2012 y 2015, Morcillo et al 2011, Martínez-Hervas et al 2012, Marrachelli et al 2014), aunque dejó de comercializarse hace unos años, posiblemente debido a la complejidad de la técnica para su empleo, y a la elevada tasa de polimorfismos que en la práctica no permitía analizar.

La secuenciación de nueva generación realizada en este trabajo se ha llevado a cabo mediante la plataforma Miseq de Illumina. En nuestro laboratorio, se ha desarrollado un protocolo para la obtención de las librerías basado en 2 PCRs en multiplex consecutivas, una es específica de la región del ADN que se quiere estudiar, y la siguiente PCR toma como molde a esa primera, para incorporarle los índices identificatorios de cada muestra. Para realizar estas amplificaciones, se han utilizado las secuencias (adaptadores) necesarias para la posterior secuenciación con la metodología Illumina. En la primer PCR (PCR1) se han usado oligos específicos del polimorfismo de interés, a los que se les han incorporado las secuencias adaptadores adecuadas, mientras que en la segunda PCR (PCR2), se han utilizado oligos de amplificación complementarios a la PCR1, que llevan los índices, en la ubicación

correspondiente para que funcionen finalmente en la secuenciación. A su vez, se han secuenciado en una misma carrera hasta 4855 muestras, a pesar de que en Illumina sólo dan la posibilidad de juntar 384 muestras según los índices comercializados por ellos. Esto se ha logrado mediante la incorporación de diferentes oligonucleótidos de 8 pares de bases, que actúan como índices, en los oligos de amplificación de la PCR2, para cada muestra. Todo este procedimiento está descrito detalladamente en la metodología. El fin de instaurar el uso de este protocolo en nuestro laboratorio, ha sido el de abaratar los costes, debido a que los kits comerciales para realizar este trabajo resultan muy caros. Nuestro procedimiento permite procesar una gran cantidad de muestras y amplicones a la vez, ahorrando también en la cantidad de ADN de cada muestra que se necesita, al ser amplificaciones en multiplex.

Limitaciones y perspectivas futuras de este trabajo

La técnica utilizada para la fase preliminar de este estudio, no ha funcionado completamente como se esperaba. De los 96 polimorfismos en 29 genes que se habían seleccionado para genotipar, sólo se han podido estudiar 66. En algunos genes, sólo hemos obtenido los genotipos de 1 variante, como es el caso de UQCRB y COX5A. En ambos, se han obtenido unas asociaciones relevantes, a pesar de ser polimorfismos que no se encuentran en las regiones codificantes del gen. Las variantes restantes de cada gen que se planeaban estudiar, y que han fallado, pertenecían la mayoría a estas regiones, con efectos funcionales más claros sobre los genes. Si hubiéramos podido analizarlas, y hubiéramos encontrado asociaciones con ellas, se podría haber hallado al SNP que podría estar afectando a la enfermedad de forma más directa. El hecho de que no se hayan genotipado la mayoría de los SNPs de un mismo gen, confirma que esto se debe a problemas atribuibles en mayor medida a la técnica SNPlex, que no ha funcionado correctamente para ciertas regiones del ADN. Por otro lado, el elevado coste de estos estudios hacía que no pudieran incluirse una gran cantidad de polimorfismos en los estudios.

El estudio de poblaciones generales, en los análisis de asociaciones entre variantes y enfermedades, tiene ciertas características a tener en cuenta. Al estudiar este tipo de

muestras, existe un amplio rango de fenotipos, pertenecientes a los distintos desórdenes que se pueden encontrar en una población. Por ejemplo, en el caso de la DMT2, estarán representados en ella los individuos sanos, prediabéticos (los que tienen el metabolismo de la glucemia alterado) y, finalmente, los pacientes diabéticos. Al realizar los análisis de asociación se comparan diabéticos y controles, de manera que dentro de los pacientes sanos, pueden incluirse pacientes con alteraciones glucémicas, o a los prediabéticos. Además, parte de los controles pueden desarrollar DMT2 en los años siguientes al estudio realizado. Por otro lado, en poblaciones generales, se tiene una proporción baja de individuos con determinadas patologías, como es el caso de la DMT2, por lo que para poder encontrar un efecto de un polimorfismo sobre una enfermedad, el tamaño muestral debe ser muy grande. Estos hechos hacen más difícil el hallazgo de asociaciones entre las variantes y las enfermedades. Nuestro trabajo es un estudio poblacional, que incluye poblaciones generales, por lo que tiene estas limitaciones. Al estudiar poblaciones más específicas, como es el caso de las subpoblaciones obtenidas según género y obesidad, es cuando hemos hallado las asociaciones que se han confirmado.

Las poblaciones estudiadas en la fase preliminar y en la de replicación, a pesar de ser poblaciones generales españolas, son muestras de distintas regiones, como se ha detallado en esta discusión. En España, existen ciertas diferencias en factores genéticos, ambientales y estilos de vida a lo largo de su territorio, que influyen en el desarrollo de la DMT2. Futuros estudios sobre parámetros ambientales (nutrición, nivel de ejercicio, etc.) que desarrollaremos en breve con estas muestras, pueden ayudar a entender las diferencias encontradas entre poblaciones.

Considerando ahora las frecuencias de las variantes asociadas, si éstas son muy bajas, se debería aumentar mucho el tamaño muestral para validar una asociación. Este efecto es más evidente cuando el SNP se asocia en un modelo recesivo, como es el caso de rs1022580 y rs2746462 con PAS, en donde sólo 3 individuos presentaban el genotipo que se había asociado significativamente. Las asociaciones replicadas y validadas en la población total, de los SNPs rs1136224, rs2410718 y

rs10742816, no han sido más significativas al aumentar el tamaño muestral, lo que se explica por la baja frecuencia del alelo asociado. Este efecto es más pronunciado para el último de ellos, en el que el genotipo asociado representa un 0.7% de los individuos no diabéticos y un 1.6% de los diabéticos, de la población de obesos.

Las variantes comunes relacionadas con la diabetes tipo 2, muestran en general 3 características (Kato et al 2013):

-El impacto genético de una variante individual es generalmente modesto, la mayoría de los odd ratios hallados para ellas están en el rango de 1.06 a 1.20.

-La mayoría de los loci identificados no están dentro de la secuencia de los principales genes candidatos y muchos se localizan en regiones intergénicas próximas al gen estudiado.

-A pesar de que el número de loci asociados son limitados, se cree que existe una especificidad poblacional en las asociaciones con la DMT2, debido a que algunas variantes parecen ejercer efectos genéticos más pronunciados en grupos étnicos específicos.

En nuestro trabajo se han obtenido resultados compatibles con estas afirmaciones. Los OR en general han sido cercanos a 1 en las asociaciones con polimorfismos más frecuentes, salvo en las asociaciones donde los individuos con el alelo asociado eran escasos en nuestra muestra (SNPs de MAFs muy bajos), lo que necesitaría una confirmación en un análisis con mayor tamaño muestral. Esta característica es más visible al aumentar el tamaño muestral, en el análisis de la población total estudiada.

Los hallazgos obtenidos en este trabajo apoyan la hipótesis de que existe una relación entre la cadena respiratoria mitocondrial y el desarrollo de la DMT2, si bien, nuestros datos no han permitido identificar variantes con un fuerte efecto en el riesgo, al analizarse en su mayoría polimorfismos frecuentes. Finalmente, hay que remarcar que las asociaciones encontradas se centran en grupos concretos de la población, identificados por el género o por la presencia/ausencia de obesidad.

Con la nueva metodología disponible en nuestro laboratorio, tenemos la oportunidad de ampliar el estudio a un gran número de variantes y genes de la cadena mitocondrial, ya sean nucleares o mitocondriales, en una población española con un tamaño muestral que podría aumentarse con respecto al de este trabajo.

7. Conclusiones

- 1- Hemos identificado varios genes de la CMR asociados con la diabetes mellitus tipo 2, obesidad, hipertensión, y parámetros relacionados, en la población general española.
- 2- El estudio de los grupos poblacionales definidos por género y por la presencia de obesidad, ha permitido identificar asociaciones relevantes y específicas de cada subgrupo.
- 3- Las variantes genéticas asociadas con DMT2 y parámetros relacionados, en la población estudiada en este trabajo son: rs10504961 (UQCRB), rs10742816 (NDUFS3), rs1136224 (NDUFS2), rs2410718 (COX7C) y el rs11205591 (NDUFS5), reforzando el hecho de que esos polimorfismos influyen en el desarrollo de la DMT2.
- 4- Nuestros resultados muestran que diferentes alteraciones de la CMR pueden estar implicadas en el desarrollo de alteraciones metabólicas, incluyendo obesidad, DMT2, y este efecto es influenciado por otros factores.
- 5- Son necesarios más análisis para confirmar estos hallazgos mediante el estudio de poblaciones más amplias, análisis de factores potencialmente moduladores (edad, sexo, BMI, ejercicio, etc.), análisis de la secuencia de los genes en la población, y estudios funcionales.

8. Bibliografía

- Abdul-Ghani MA, DeFronzo RA. Mitochondrial dysfunction, insulin resistance, and type 2 diabetes mellitus. *Curr Diab Rep.* 2008; 8(3):173-8.
- Adlan MA, Bondugulapati LN, Premawardhana LD. Glucose intolerance and diabetes mellitus in endocrine disorders - two case reports and a review. *Curr Diabetes Rev.* 2010; 6(5):266-73.
- Ahlqvist E, Ahluwalia TS, Groop L. Genetics of type 2 diabetes. *Clin Chem.* 2011; 57(2):241-54.
- Akude E, Zhrebetskaya E, Chowdhury SK, Smith DR, Dobrowsky RT, Ferryhough P. Diminished superoxide generation is associated with respiratory chain dysfunction and changes in the mitochondrial proteome of sensory neurons from diabetic rats. *Diabetes.* 2011; 60(1):288-97.
- Ali O. Genetics of type 2 diabetes. *World J Diabetes.* 2013; 4(4):114-23.
- Alkhalaf A, Landman GW, van Hateren KJ, Groenier KH, Mooyaart AL, De Heer E, Gans RO, Navis GJ, Bakker SJ, Kleefstra N, Bilo HJ. Sex specific association between carnosinase gene CNDP1 and cardiovascular mortality in patients with type 2 diabetes (ZODIAC-22). *J Nephrol.* 2015; 28(2):201-7.
- Allison DB, Kaprio J, Korkeila M, Koskenvuo M, Neale MC, Hayakawa K. The heritability of body mass index among an international sample of monozygotic twins reared apart. *Int J Obes Relat Metab Disord* 1996; 20: 501 – 506.
- American Diabetes Association (ADA). Classification and Diagnosis of Diabetes. *Diabetes Care.* 2016; 39 Suppl 1:S13-22.
- American Diabetes Association (ADA). Diagnosis and classification of diabetes mellitus. *Diabetes Care.* 2012; 35 Suppl 1:S64-71.
- American Diabetes Association (ADA). Diagnosis and classification of diabetes mellitus. *Diabetes Care.* 2014 Jan;37 (Suppl 1):S81-90.
- American Diabetes Association (ADA). Standards of Medical Care in Diabetes 2015. *Diabetes Care* 2015;38(Suppl. 1):S1-S94.
- American Diabetes Association (ADA). Standards of medical care in diabetes--2011. *Diabetes Care.* 2011 Jan; 34 (Suppl 1):S11-61.
- Andreasen CH, Stender-Petersen KL, Mogensen MS, Torekov SS, Wegner L, Andersen G, Nielsen AL, Albrechtsen A, Borch-Johnsen K, Rasmussen SS, Clausen JO, Sandbaek A, Lauritzen T, Hansen L, Jørgensen T, Pedersen O, Hansen T. Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes* 2008; 57(1):95-101.
- Ban M, Elson J, Walton A, Turnbull D, Compston A, Chinnery P, Sawcer S. Investigation of the role of mitochondrial ADN in multiple sclerosis susceptibility. *PLoS One.* 2008; 3(8):e2891.
- Barrett TG. Mitochondrial diabetes, DIDMOAD and other inherited diabetes syndromes. *Best Pract Res Clin Endocrinol Metab.* 2001; 15(3): 325-43.
- Basile KJ, Johnson ME, Xia Q, Grant SF. Genetic susceptibility to type 2 diabetes and obesity: follow-up of findings from genome-wide association studies. *Int J Endocrinol.* 2014; 2014:769-671.
- Basterra-Gortari FJ, Martínez-González MA. Comparación de la prevalencia de diabetes mellitus entre comunidades autónomas. *Med Clin (Barc).* 2007; 129(18):716-9.
- Batkai S, Thum T. MicroRNAs in hypertension: mechanisms and therapeutic targets. *Curr Hypertens Rep* 2012; 14: 79–87.
- Becker ML, Pearson ER, Tkáč I. Pharmacogenetics of oral antidiabetic drugs. *Int J Endocrinol.* 2013;2013:686315.

- Billings LK, Florez JC. The genetics of type 2 diabetes: what have we learned from GWAS? *Annals of the New York Academy of Sciences*. 2010; 1212:59-77.
- Bonnefond A, Froguel P, Vaxillaire M. The emerging genetics of type 2 diabetes. *Trends in molecular medicine*. 2010; 16(9):407-16.
- Bowsher CG, Tobin AK. Compartmentation of metabolism within mitochondria and plastids. *Journal of Experimental Botany* 2001; 52: 513–527.
- Boyle KE, Hwang H, Janssen RC, DeVente JM, Barbour LA, Hernandez TL, Mandarino LJ, Lappas M, Friedman JE. Gestational diabetes is characterized by reduced mitochondrial protein expression and altered calcium signaling proteins in skeletal muscle. *PLoS One*. 2014; 9(9):e106872.
- Bratic A, Larsson NG. The role of mitochondria in aging. *J Clin Invest*. 2013 1;123(3):951-7.
- Brito EC, Lyssenko V, Renström F, Berglund G, Nilsson PM, Groop L, Franks PW. Previously associated type 2 diabetes variants may interact with physical activity to modify the risk of impaired glucose regulation and type 2 diabetes: a study of 16,003 Swedish adults. *Diabetes* 2009; 58(6): 1411-8.
- Brunetti A, Chiefari E, Foti D. Recent advances in the molecular genetics of type 2 diabetes mellitus. *World J Diabetes*. 2014;5(2):128-40.
- Burant CF, Viswanathan P, Marcinak J, et al. TAK-875 versus placebo or glimepiride in type 2 diabetes mellitus: a phase 2, randomised, double-blind, placebo-controlled trial. *Lancet*. 2012; 379:1403–1411.
- Calvanese V, Lara E, Kahn A, Fraga MF. The role of epigenetics in aging and age-related diseases. *Ageing Res Rev*. 2009; 8(4):268-76.
- Carrozzo R, Dionisi-Vici C, Steuerwald U, Luciola S, Deodato, F, Di Giandomenico S, Bertini E, Franke B, Kluijtmans LA, Meschini MC, Rizzo C, Piemonte F, Rodenburg R, Santer R, Santorelli FM, van Rooij A, Vermunt-de Koning D, Morava E, Wevers RA. SUCLA2 mutations are associated with mild methylmalonic aciduria, Leigh-like encephalomyopathy, dystonia and deafness *Brain* 2007; 130: 862–874.
- Chao EC, Henry RR. SGLT2 inhibition – a novel strategy for diabetes treatment. *Nat Rev Drug Discov*. 2010;9:551–559.
- Cho YM, Shin HD, Park BL, Kim JH, Park KS, Kim SY, Lee HK. Association between polymorphisms in the nuclear respiratory factor 1 gene and type 2 diabetes mellitus in the Korean population. *Diabetologia*. 2005; 48(10):2033-8.
- Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat.Rev.Genet*. 2010; 11:415–425.
- Colosia AD, Palencia R, Khan S. Prevalence of hypertension and obesity in patients with type 2 diabetes mellitus in observational studies: a systematic literature review. *Diabetes Metab Syndr Obes*. 2013; 17(6):327-38.
- Cornelis MC, Hu FB. Gene-environment interactions in the development of type 2 diabetes: recent progress and continuing challenges. *Annu Rev Nutr*. 2012; 32:245-59.
- Cornelis MC, Qi L, Kraft P, Hu FB. TCF7L2, dietary carbohydrate, and risk of type 2 diabetes in US women. *Am J Clin Nutr* 2009; 89:1256–1262.
- Cornell S. Continual evolution of type 2 diabetes: an update on pathophysiology and emerging treatment options. *Ther Clin Risk Manag*. 2015; 16(11):621-32.
- Crunkhorn S, Dearie F, Mantzoros C, Gami H, da Silva WS, Espinoza D, Faucette R, Barry K, Bianco AC, Patti ME. Peroxisome proliferator activator receptor γ coactivator-1 expression is reduced in obesity: potential pathogenic role of saturated fatty acids and p38 mitogen-activated protein kinase activation. *J Biol Chem* 2007; 282:15439–15450.

- Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, Lin JK, Farzadfar F, Khang YH, Stevens GA, Rao M, Ali MK, Riley LM, Robinson CA, Ezzati M; Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Blood Glucose). National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *Lancet*. 2011; 378(9785):31-40.
- Delettre C, Lenaers G, Griffoin JM, Gigarel N, Lorenzo C, Belenguer P, Pelloquin L, Grosgeorge J, Turc-Carel C, Perret E, Astarie-Dequeker C, Lasquellec L, Arnaud B, Ducommun B, Kaplan J, Hamel CP. Nuclear gene OPA1, encoding a mitochondrial dynamin-related protein, is mutated in dominant optic atrophy. *Nature Genetics* 2000; 26: 207–210.
- Dikalov SI, Dikalova AE. Contribution of mitochondrial oxidative stress to hypertension. *Curr Opin Nephrol Hypertens*. 2016; 25(2):73-80.
- Douce R, Neuburger M. Biochemical dissection of photorespiration. *Current Opinion in Plant Biology* 1999;2: 214–222.
- Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL, Leach RJ, O’Connell P, Stern MP. Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromosome 10q in Mexican Americans. *Am J Hum Genet* 1999; 64:1127-1140.
- Ek J, Andersen G, Urhammer SA, Gaede PH, Drivsholm T, Borch-Johnsen K, Hansen T, Pedersen O. Mutation analysis of peroxisome proliferator-activated receptor-gamma coactivator-1 (PGC-1) and relationships of identified amino acid polymorphisms to Type II diabetes mellitus. *Diabetologia* 2001; 44(12):2220-6.
- Ellard S & Colclough K. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha (HNF1A) and 4 alpha (HNF4A) in maturity-onset diabetes of the young. *Hum Mutat* 2006; 27: 854–869.
- Eze IC, Schaffner E, Fischer E, Schikowski T, Adam M, Imboden M, Tsai M, Carballo D, von Eckardstein A, Künzli N, Schindler C, Probst-Hensch N. Long-term air pollution exposure and diabetes in a population-based Swiss cohort. *Environ Int*. 2014; 70:95-105.
- Farooqi IS, Keogh JM, Yeo GS, Lank EJ, Cheetham T, O’Rahilly S. Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N. Engl. J. Med*. 2003, 348: 1085–1095.
- Ferrannini E, Buzzigoli G, Bonadonna R, Giorico MA, Oleggini M, Graziadei L, Pedrinelli R, Brandi L, Bevilacqua S. Insulin resistance in essential hypertension. *N Engl J Med*. 1987; 317(6): 350-7.
- Ferrannini E, Cushman WC. Diabetes and hypertension: the bad companions. *Lancet*. 2012; 380(9841): 601-10.
- Fioretto P, Giaccari A, Sesti G. Efficacy and safety of dapagliflozin, a sodium glucose cotransporter 2 (SGLT2) inhibitor, in diabetes mellitus. *Cardiovasc Diabetol*. 2015; 14:142.
- Flannick J, Thorleifsson G, Beer NL, Jacobs SB, Grarup N, Burtt NP, Mahajan A, Fuchsberger C, Atzmon G, Benediktsson R, Blangero J, Bowden DW, Brandslund I, Brosnan J, Burslem F, Chambers J, Cho YS, Christensen C, Douglas DA, Duggirala R, Dymek Z, Farjoun Y, Fennell T, Fontanillas P, Forsén T, Gabriel S, Glaser B, Gudbjartsson DF, Hanis C, Hansen T, Hreidarsson AB, Hveem K, Ingelsson E, Isomaa B, Johansson S, Jørgensen T, Jørgensen ME, Kathiresan S, Kong A, Kooner J, Kravic J, Laakso M, Lee JY, Lind L, Lindgren CM, Linneberg A, Masson G, Meitinger T, Mohlke KL, Molven A, Morris AP, Potluri S, Rauramaa R, Ribel-Madsen R, Richard AM, Rolph T, Salomaa V, Segrè AV, Skärstrand H, Steinthorsdottir V, Stringham HM, Sulem P, Tai ES, Teo YY, Teslovich T,

- Thorsteinsdottir U, Trimmer JK, Tuomi T, Tuomilehto J, Vaziri-Sani F, Voight BF, Wilson JG, Boehnke M, McCarthy MI, Njølstad PR, Pedersen O; Go-TD Consortium; TD-GENES Consortium, Groop L, Cox DR, Stefansson K, Altshuler D. Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat Genet.* 2014; 46(4):357-63.
- Florez JC, Hirschhorn J, Altshuler D. The inherited basis of diabetes mellitus: implications for the genetic analysis of complex traits. *Annu Rev Genomics Hum Genet* 2003; 4: 257-291.
 - Florez JC. Clinical review: the genetics of type 2 diabetes: a realistic appraisal in 2008. *The Journal of clinical endocrinology and metabolism.* 2008; 93(12):4633-42.
 - Fontanesi L, Galimberti G, Calò DG, Fronza R, Martelli PL, Scotti E, Colombo M, Schiavo G, Casadio R, Buttazzoni L, Russo V. Identification and association analysis of several hundred single nucleotide polymorphisms within candidate genes for back fat thickness in Italian Large White pigs using a selective genotyping approach. *J Anim Sci.* 2012; 90(8):2450-64.
 - Fosslien E. Mitochondrial medicine--molecular pathology of defective oxidative phosphorylation. *Ann Clin Lab Sci.* 2001; 31(1):25-67.
 - Franks PW, Pearson E, Florez JC. Gene-environment and gene-treatment interactions in type 2 diabetes: progress, pitfalls, and prospects. *Diabetes Care* 2013; 36(5):1413-21.
 - Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JR, Elliott KS, Lango H, Rayner NW, Shields B, Harries LW, Barrett JC, Ellard S, Groves CJ, Knight B, Patch AM, Ness AR, Ebrahim S, Lawlor DA, Ring SM, Ben-Shlomo Y, Jarvelin MR, Sovio U, Bennett AJ, Melzer D, Ferrucci L, Loos RJ, Barroso I, Wareham NJ, Karpe F, Owen KR, Cardon LR, Walker M, Hitman GA, Palmer CN, Doney AS, Morris AD, Smith GD, Hattersley AT, McCarthy MI. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science.* 2007; 316(5826):889-94.
 - Fujiki K, Kano F, Shiota K, Murata M. Expression of the peroxisome proliferator activated receptor gamma gene is repressed by ADN methylation in visceral adipose tissue of mouse models of diabetes. *BMC Biol.* 2009; 10;7:38.
 - Galan-Chilet I, Grau-Perez M, De Marco G, Guallar E, Martin-Escudero JC, Dominguez-Lucas A, Gonzalez-Manzano I, Lopez-Izquierdo R, Briongos-Figuero LS, Redon J, Chaves FJ, Tellez-Plaza M. A gene-environment interaction analysis of plasma selenium with prevalent and incident diabetes: The Hortega study. *Redox Biol.* 2017; 12:798-805.
 - Gerbitz K, Gempel K & Brdiczka D. Mitochondria and diabetes. *Diabetes* 1996; 45: 113-126.
 - Gilbert ER, Liu D. Epigenetics: the missing link to understanding β -cell dysfunction in the pathogenesis of type 2 diabetes. *Epigenetics.* 2012;7(8):841-52.
 - GlaxoSmithKline. A study in type 2 diabetics of single and multiple doses of orally administered GSK1292263 to investigate the safety, tolerability, pharmacokinetics and pharmacodynamics. Available from: <http://clinicaltrials.gov/ct2/show/NCT01119846?term=gsk1292263&rank=1>. Accessed July 17, 2015.
 - Gonçalves VF, Zai CC1, Tiwari AK, Brandl EJ, Derkach A, Meltzer HY, Lieberman JA, Müller DJ, Sun L, Kennedy JL. A hypothesis-driven association study of 28 nuclear-encoded mitochondrial genes with antipsychotic-induced weight gain in schizophrenia. *Neuropsychopharmacology.* 2014; 39(6):1347-54.

- Gray J, Yeo GS, Cox JJ, Morton J, Adlam AL, Keogh JM, Yanovski JA, El Gharbawy A, Han JC, Tung YC, Hodges JR, Raymond FL, O'rahilly S, Farooqi IS. Hyperphagia, severe obesity, impaired cognitive function, and hyperactivity associated with functional loss of one copy of the brain-derived neurotrophic factor (BDNF) gene. *Diabetes* 2006; 55, 3366-3371.
- Groop LC, Bottazzo GF, Doniach D. Islet cell antibodies identify latent type 1 diabetes in patients aged 35–75 years at diagnosis. *Diabetes* 1986; 17:1214–1219.
- Grossman E, Messerli FH. Hypertension and diabetes. *Adv Cardiol.* 2008; 45: 82-106.
- Guay C, Roggli E, Nesca V, Jacovetti C, Regazzi R. Diabetes mellitus, a microRNA-related disease? *Transl Res* 2011; 157(4): 253-64.
- Gutiérrez-Fisac JL, Guallar-Castillón P, León-Muñoz LM, Graciani A, Banegas JR, Rodríguez-Artalejo F. Prevalence of general and abdominal obesity in the adult population of Spain, 2008-2010: the ENRICA study. *Obes Rev.* 2012; 13(4):388-92.
- Hani EH1, Boutin P, Durand E, Inoue H, Permutt MA, Velho G, Froguel P. Missense mutations in the pancreatic islet beta cell inwardly rectifying K⁺ channel gene (KIR6.2/BIR): a meta-analysis suggests a role in the polygenic basis of Type II diabetes mellitus in Caucasians. *Diabetologia* 1998; 41(12):1511-5.
- Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, Wapelhorst B, Spielman RS, Gogolin-Ewens KJ, Shepard JM, Williams SR, Risch N, Hinds D, Iwasaki N, Ogata M, Omori Y, Petzold C, Rietzch H, Schröder HE, Schulze J, Cox NJ, Menzel S, Boriraj VV, Chen X, Lim LR, Lindner T, Mereu LE, Wang YQ, Xiang K, Yamagata K, Yang Y, Bell GI. A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat Genet* 1996; 13: 161-166.
- Hara K, Tobe K, Okada T, Kadowaki H, Akanuma Y, Ito C, Kimura S, Kadowaki T. A genetic variation in the PGC-1 gene could confer insulin resistance and susceptibility to Type II diabetes. *Diabetologia* 2002; 45(5):740-3.
- Hegele RA, Zinman B, Hanley AJ, Harris SB, Barrett PH, Cao H. Genes, environment and Oji-Cree type 2 diabetes. *Clin Biochem.* 2003; 36(3):163-70.
- Holstein AI, Plaschke A, Ptak M, Egberts EH, El-Din J, Brockmüller J, Kirchheiner J. Association between CYP2C9 slow metabolizer genotypes and severe hypoglycaemia on medication with sulphonylurea hypoglycaemic agents. *Br J Clin Pharmacol.* 2005;60(1):103-6.
- Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, Lindner TH, Mashima H, Schwarz PE, del Bosque-Plata L, Horikawa Y, Oda Y, Yoshiuchi I, Colilla S, Polonsky KS, Wei S, Concannon P, Iwasaki N, Schulze J, Baier LJ, Bogardus C, Groop L, Boerwinkle E, Hanis CL, Bell GI. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat Genet* 2000; 26:163-175.
- Imam K. Clinical features, diagnostic criteria and pathogenesis of diabetes mellitus. *Adv Exp Med Biol.* 2012; 771:340-55.
- Imamura M, Maeda S. Genetics of type 2 diabetes: the GWAS era and future perspectives. *Endocr J.* 2011; 58(9): 723-39.
- InterAct Consortium. The link between family history and risk of type 2 diabetes is not explained by anthropometric, lifestyle or genetic risk factors: the EPIC-InterAct study. *Diabetologia.* 2013; 56(1):60-9.
- International Consortium for Blood Pressure Genome-Wide Association Studies, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, Smith AV, Tobin MD, Verwoert GC, Hwang SJ, Pihur V, Vollenweider P, O'Reilly PF, Amin

N, Bragg-Gresham JL, Teumer A, Glazer NL, Launer L, Zhao JH, Aulchenko Y, Heath S, Söber S, Parsa A, Luan J, Arora P, Dehghan A, Zhang F, Lucas G, Hicks AA, Jackson AU, Peden JF, Tanaka T, Wild SH, Rudan I, Igl W, Milaneschi Y, Parker AN, Fava C, Chambers JC, Fox ER, Kumari M, Go MJ, van der Harst P, Kao WH, Sjögren M, Vinay DG, Alexander M, Tabara Y, Shaw-Hawkins S, Whincup PH, Liu Y, Shi G, Kuusisto J, Tayo B, Seielstad M, Sim X, Nguyen KD, Lehtimäki T, Matullo G, Wu Y, Gaunt TR, Onland-Moret NC, Cooper MN, Platou CG, Org E, Hardy R, Dahgam S, Palmen J, Vitart V, Braund PS, Kuznetsova T, Uiterwaal CS, Adeyemo A, Palmas W, Campbell H, Ludwig B, Tomaszewski M, Tzoulaki I, Palmer ND; CARDIoGRAM consortium; CKDGen Consortium; KidneyGen Consortium; EchoGen consortium; CHARGE-HF consortium, Aspelund T, Garcia M, Chang YP, O'Connell JR, Steinle NI, Grobbee DE, Arking DE, Kardia SL, Morrison AC, Hernandez D, Najjar S, McArdle WL, Hadley D, Brown MJ, Connell JM, Hingorani AD, Day IN, Lawlor DA, Beilby JP, Lawrence RW, Clarke R, Hopewell JC, Ongen H, Dreisbach AW, Li Y, Young JH, Bis JC, Kähönen M, Viikari J, Adair LS, Lee NR, Chen MH, Olden M, Pattaro C, Bolton JA, Köttgen A, Bergmann S, Mooser V, Chaturvedi N, Frayling TM, Islam M, Jafar TH, Erdmann J, Kulkarni SR, Bornstein SR, Grässler J, Groop L, Voight BF, Kettunen J, Howard P, Taylor A, Gurrera S, Ricceri F, Emilsson V, Plump A, Barroso I, Khaw KT, Weder AB, Hunt SC, Sun YV, Bergman RN, Collins FS, Bonnycastle LL, Scott LJ, Stringham HM, Peltonen L, Perola M, Vartiainen E, Brand SM, Staessen JA, Wang TJ, Burton PR, Soler Artigas M, Dong Y, Snieder H, Wang X, Zhu H, Lohman KK, Rudock ME, Heckbert SR, Smith NL, Wiggins KL, Doumatey A, Shriner D, Veldre G, Viigimaa M, Kinra S, Prabhakaran D, Tripathy V, Langefeld CD, Rosengren A, Thelle DS, Corsi AM, Singleton A, Forrester T, Hilton G, McKenzie CA, Salako T, Iwai N, Kita Y, Ogihara T, Ohkubo T, Okamura T, Ueshima H, Umemura S, Eyheramendy S, Meitinger T, Wichmann HE, Cho YS, Kim HL, Lee JY, Scott J, Sehmi JS, Zhang W, Hedblad B, Nilsson P, Smith GD, Wong A, Narisu N, Stančáková A, Raffel LJ, Yao J, Kathiresan S, O'Donnell CJ, Schwartz SM, Ikram MA, Longstreth WT Jr, Mosley TH, Seshadri S, Shrine NR, Wain LV, Morken MA, Swift AJ, Laitinen J, Prokopenko I, Zitting P, Cooper JA, Humphries SE, Danesh J, Rasheed A, Goel A, Hamsten A, Watkins H, Bakker SJ, van Gilst WH, Janipalli CS, Mani KR, Yajnik CS, Hofman A, Mattace-Raso FU, Oostra BA, Demirkan A, Isaacs A, Rivadeneira F, Lakatta EG, Orru M, Scuteri A, Ala-Korpela M, Kangas AJ, Lytikäinen LP, Soininen P, Tukiainen T, Würtz P, Ong RT, Dörr M, Kroemer HK, Völker U, Völzke H, Galan P, Hercberg S, Lathrop M, Zelenika D, Deloukas P, Mangino M, Spector TD, Zhai G, Meschia JF, Nalls MA, Sharma P, Terzic J, Kumar MV, Denniff M, Zukowska-Szczechowska E, Wagenknecht LE, Fowkes FG, Charchar FJ, Schwarz PE, Hayward C, Guo X, Rotimi C, Bots ML, Brand E, Samani NJ, Polasek O, Talmud PJ, Nyberg F, Kuh D, Laan M, Hveem K, Palmer LJ, van der Schouw YT, Casas JP, Mohlke KL, Vineis P, Raitakari O, Ganesh SK, Wong TY, Tai ES, Cooper RS, Laakso M, Rao DC, Harris TB, Morris RW, Dominiczak AF, Kivimäki M, Marmot MG, Miki T, Saleheen D, Chandak GR, Coresh J, Navis G, Salomaa V, Han BG, Zhu X, Kooner JS, Melander O, Ridker PM, Bandinelli S, Gyllenstein UB, Wright AF, Wilson JF, Ferrucci L, Farrall M, Tuomilehto J, Pramstaller PP, Elosua R, Soranzo N, Sijbrands EJ, Altshuler D, Loos RJ, Shuldiner AR, Gieger C, Meneton P, Uitterlinden AG, Wareham NJ, Gudnason V, Rotter JI, Rettig R, Uda M, Strachan DP, Witteman JC, Hartikainen AL, Beckmann JS, Boerwinkle E, Vasani RS, Boehnke M, Larson MG, Jarvelin MR, Psaty BM, Abecasis GR, Chakravarti A, Elliott P, van Duijn CM, Newton-Cheh C, Levy D, Caulfield MJ, Johnson T. Genetic variants in novel

- pathways influence blood pressure and cardiovascular disease risk. *Nature*. 2011; 478(7367): 103-9.
- Jackson RS, Creemers JW, Ohagi S, Raffin-Sanson ML, Sanders L, Montague CT, Hutton JC, O'Rahilly S.. Obesity and impaired prohormone processing associated with mutations in the human prohormoneconvertase 1 gene. *Nat. Genet.* 1997; 16: 303–306.
 - Jamaluddin JL1, Huri HZ, Vethakkan SR, Mustafa N. Pancreatic gene variants potentially associated with dipeptidyl peptidase-4 inhibitor treatment response in Type 2 diabetes. *Pharmacogenomics*. 2014; 15(2):235-49.
 - Jitrapakdee S, Wutthisathapornchai A, Wallace JC, MacDonald MJ. Regulation of insulin secretion: role of mitochondrial signalling. *Diabetologia*. 2010; 53(6):1019-32.
 - Johannsen DL, Ravussin E: The role of mitochondria in health and disease. In *Curr Opin Pharmacol England*, 2009; 780-786.
 - Jones A. Does the plant mitochondrion integrate cellular stress and regulate programmed cell death? *Trends in Plant Science* 2000; 5:225–230.
 - Kadenbach, B (Ed.). *Mitochondrial Oxidative Phosphorylation. Nuclear-Encoded Genes, Enzyme Regulation, and Pathophysiology. Advances in Experimental Medicine and Biology*, 2012; Vol. 748.
 - Kadowaki T, Kadowaki H, Mori Y, Tobe K, Sakuta R, Suzuki Y, Tanabe Y, Sakura H, Awata T, Goto Y, et al. A subtype of diabetes mellitus associated with a mutation of mitochondrial ADN. *New England Journal of Medicine* 1994; 330: 962-968.
 - Kalendar R, Lee D, Schulman AH. FastPCR software for PCR, in silico PCR, and oligonucleotide assembly and analysis. *Methods Mol Biol*. 2014; 1116:271-302.
 - Kameoka K, Isotani H, Tanaka K, Azukari K, Fujimura Y, Shiota Y, Sasaki E, Majima M, Furukawa K, Haginomori S, Kitaoka H, Ohsawa N. Novel mitochondrial ADN mutation in tRNA(Lys) (8296A-G) associated with diabetes. *Biochemical and Biophysical Research Communications* 1998; 17: 523-527.
 - Karani S, Vimaleswaran KS, Loos RJ. Progress in the genetics of common obesity and type 2 diabetes. *Expert Rev Mol Med*. 2010; 26; 12:e7.
 - Kato N. Insights into the genetic basis of type 2 diabetes. *J Diabetes Investig*. 2013; 4(3):233-44.
 - Kaukonen J, Juselius JK, Tiranti V, Kyttala A, Zeviani M, Comi GP, Keranen S, Peltonen L, Suomalainen A. Role of adenine nucleotide translocator 1 in mtADN maintenance. *Science* 2000; 289: 782–785.
 - Kim JA, Wei Y, Sowers JR. Role of mitochondrial dysfunction in insulin resistance. *Circ Res*. 2008; 102(4):401-14.
 - Kleinberger JW, Pollin TI. Personalized medicine in diabetes mellitus: current opportunities and future prospects. *Ann N Y Acad Sci*. 2015;1346(1):45-56.
 - Kobayashi T, Tamemoto K, Nakanishi K, Kato N, Okubo M, Kajio H, Sugimoto T, Murase T, Kosaka K. Immunogenetic and clinical characterization of slowly progressive IDDM. *Diabetes Care* 1993; 16:780–788.
 - Kong L, Zhu J, Han W, Jiang X, Xu M, Zhao Y, Dong Q, Pang Z, Guan Q, Gao L, Zhao J, Zhao L.
 - Krude H, Biebermann H, Luck W, Horn R, Brabant G, Grüters A. Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nat. genet.* 1998; 19: 155–157.
 - Kuo CC, Moon K, Thayer KA, Navas-Acien A. Environmental chemicals and type 2 diabetes: an updated systematic review of the epidemiologic evidence. *Curr Diab Rep*. 2013; 13(6):831-49.

- Langenberg C, Sharp SJ, Franks PW, Scott RA, Deloukas P, Forouhi NG, Froguel P, Groop LC, Hansen T, Palla L, Pedersen O, Schulze MB, Tormo MJ, Wheeler E, Agnoli C, Arriola L, Barricarte A, Boeing H, Clarke GM, Clavel-Chapelon F, Duell EJ, Fagherazzi G, Kaaks R, Kerrison ND, Key TJ, Khaw KT, Kröger J, Lajous M, Morris AP, Navarro C, Nilsson PM, Overvad K, Palli D, Panico S, Quirós JR, Rolandsson O, Sacerdote C, Sánchez MJ, Slimani N, Spijkerman AM, Tumino R, van der A DL, van der Schouw YT, Barroso I, McCarthy MI, Riboli E, Wareham NJ. Gene-lifestyle interaction and type 2 diabetes: the EPIC interact case-cohort study. *PLoS medicine*. 2014; 11(5):e1001647.
- Lee DH, Lee IK, Song K, Steffes M, Toscano W, Baker BA, Jacobs DR Jr. A strong dose-response relation between serum concentrations of persistent organic pollutants and diabetes: results from the National Health and Examination Survey 1999-2002. *Diabetes Care* 2006; 29:1638.
- Lee HA, Cho HM, Lee DY, Kim KC, Han HS, Kim IK. Tissue-specific upregulation of angiotensin-converting enzyme 1 in spontaneously hypertensive rats through histone code modifications. *Hypertension*. 2012; 59(3): 621-6.
- Li H, Achour I, Bastarache L, Berghout J, Gardeux V, Li J, Lee Y, Pesce L, Yang X, Ramos KS, Foster I, Denny JC, Moore JH, Lussier YA. Integrative genomics analyses unveil downstream biological effectors of disease-specific polymorphisms buried in intergenic regions. *NPJ Genom Med*. 2016; 1, pii: 16006.
- Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler Transform. *Bioinformatics* 2010; 26(5):589-95.
- Li SQ, Zhu JG, Zhang WL, et al. Signature microRNA expression profile of essential hypertension and its novel link to human cytomegalovirus infection. *Circulation* 2011; 124:175–U154.
- Lin MT, Beal MF. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* 2006; 443:787–795.
- Lin Y, Sun Z. Current views on type 2 diabetes. *J Endocrinol*. 2010 Jan; 204(1):1-11.
- Ling C, Groop L. Epigenetics: a molecular link between environmental factors and type 2 diabetes. *Diabetes* 2009; 58:2718–2725.
- Ling C, Poulsen P, Carlsson E, Ridderstråle M, Almgren P, Wojtaszewski J, Beck-Nielsen H, Groop L, Vaag A. Multiple environmental and genetic factors influence skeletal muscle PGC-1 α and PGC-1 β gene expression in twins. *J Clin Invest* 2004; 114:1518–1526.
- Ling C, Poulsen P, Simonsson S, Ronn T, Holmkvist J, Almgren P, Hagert P, Nilsson E, Mabey AG, Nilsson P, Vaag A & Groop L. Genetic and epigenetic factors are associated with expression of respiratory chain component NDUFB6 in human skeletal muscle. *J Clin Invest*. 2007; 117: 3427–3435.
- Liu C, Ying Z, Harkema J, Sun Q, Rajagopalan S. Epidemiological and experimental links between air pollution and type 2 diabetes. *Toxicol Pathol*. 2013;41(2):361-73.
- Logan DC, Knight MR. Mitochondrial and cytosolic calcium dynamics are differentially regulated in plants. *Plant Physiology* 2003; 133: 21–24.
- Logan DC. The mitochondrial compartment. *Journal of Experimental Botany* 2006; 57 (6); 1225–1243.
- López-Sobaler AM, Aparicio A, Aranceta-Bartrina J, Gil A, González-Gross M, Serra-Majem L, Varela-Moreiras G, Ortega RM. Overweight and General and Abdominal Obesity in a Representative Sample of Spanish Adults: Findings from the ANIBES Study *Biomed Res Int*. 2016; 2016: 8341487.
- López-Sobaler AM, Rodríguez-Rodríguez E, Aranceta-Bartrina J, Gil A, González-Gross M, Serra-Majem L, Varela-Moreiras G, Ortega RM. General and Abdominal

Obesity Is Related to Physical Activity, Smoking and Sleeping Behaviours and Mediated by the Educational Level: Findings from the ANIBES Study in Spain. *PLoS One*. 2016; 11(12): e0169027.

- Lu H, Yang Y, Allister EM, Wijesekara N, Wheeler MB. The identification of potential factors associated with the development of type 2 diabetes: a quantitative proteomics approach. *Mol Cell Proteomics*. 2008; 7(8):1434-51.
- Maassen J & Kadowaki T. Maternally inherited diabetes and deafness: a new diabetes subtype. *Diabetologia* 1996; 39: 375-382.
- Mackenzie S, McIntosh L. Higher plant mitochondria. *The Plant Cell* 1999;11: 571-586.
- Majamaa K, Moilanen JS, Uimonen S, Remes AM, Salmela PI, Kärppä M, Majamaa-Voltti KA, Rusanen H, Sorri M, Peuhkurinen KJ, Hassinen IE. Epidemiology of A3243G, the mutation for mitochondrial encephalomyopathy, lactic acidosis, and strokelike episodes: prevalence of the mutation in an adult population. *Am J Hum Genet*. 1998; 63(2):447-54.
- Majithia AR, Flannick J, Shahinian P, Guo M, Bray MA, Fontanillas P, Gabriel SB; GoT2D Consortium; NHGRI JHS/FHS Allelic Spectrum Project; SIGMA T2D Consortium; T2D-GENES Consortium, Rosen ED, Altshuler D. Rare variants in PPARG with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes. *Proc Natl Acad Sci U S A*. 2014; 111(36):13127-32.
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy ML, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. *Nature* 2009; 461, 747-753.
- Mansego ML, De Marco G, Alonso MP, Martínez F, Sáez GT, Escudero JC, Redón J, Chaves FJ. Polymorphisms of antioxidant enzymes, blood pressure and risk of hypertension. *J Hypertens*. 2011; 29(3): 492-500.
- Mansego ML, De Marco G, Alonso MP, Martínez F, Sáez GT, Escudero JC, Redón J, Chaves FJ. Polymorphisms of antioxidant enzymes, blood pressure and risk of hypertension. *J Hypertens*. 2011; 29(3):492-500.
- Mansego ML, De Marco G, Ivorra C, Lopez-Izquierdo R, Morcillo S, Rojo-Martínez G, González-Albert V, Martínez F, Soriguer F, Martín-Escudero JC, Redon J, Chaves FJ. The nutrigenetic influence of the interaction between dietary vitamin E and TXN and COMT gene polymorphisms on waist circumference: a case control study. *J Transl Med*. 2015; 13:286.
- Mansego ML, Martínez F, Martínez-Larrad MT, Zabena C, Rojo G, Morcillo S, Soriguer F, Martín-Escudero JC, Serrano-Ríos M, Redon J, Chaves FJ. Common variants of the liver fatty acid binding protein gene influence the risk of type 2 diabetes and insulin resistance in Spanish population. *PLoS One* 2012; 7(3):e31853.
- Manucha W, Ritchie B, Ferder L. Hypertension and insulin resistance: implications of mitochondrial dysfunction. *Curr Hypertens Rep*. 2015; 17(1):504.
- Marrachelli VG, Monleon D, Rentero P, Mansego ML, Morales JM, Galan I, Segura R, Martinez F, Martin-Escudero JC, Briongos L, Marin P, Lliso G, Chaves FJ, Redon J. Genomic and metabolomic profile associated to microalbuminuria. *PLoS One*. 2014; 9(2):e98227.
- Marti A, Ordovas J. Epigenetics lights up the obesity field. *Obes Facts*. 2011; 4(3):187-90.

- Martin SD, McGee SL. The role of mitochondria in the aetiology of insulin resistance and type 2 diabetes. *Biochim Biophys Acta*. 2013 pii: S0304-4165(13)00400-5.
- Martínez-Hervas S, Mansego ML, De Marco G, Martínez F, Alonso MP, Morcillo S, Rojo-Martínez G, Real JT, Ascaso JF, Redon J, Martín Escudero JC, Soriguer F, Chaves FJ. Polymorphisms of the UCP2 gene are associated with body fat distribution and risk of abdominal obesity in Spanish population. *Eur J Clin Invest*. 2012; 42(2):171-8.
- Martínez-Larrad MT, Fernández-Pérez C, González-Sánchez JL, López A, Fernández-Alvarez J, Riviriego J, Serrano-Ríos M; Grupo de Estudio de Atención Primaria de Segovia. Prevalencia del síndrome metabólico (criterios del ATP-III). Estudio de base poblacional en áreas rural y urbana de la provincia de Segovia. *Med Clin (Barc)*. 2005 Oct 15;125(13):481-6.
- Medici F, Hawa M, Ianari A, Pyke DA, Leslie RD. Concordance rate for type II diabetes mellitus in monozygotic twins: actuarial analysis. *Diabetologia* 1999; 42(2):146-50.
- Mena-Martin FJ, Martín-Escudero JC, Simal-Blanco F, Carretero-Ares JL, Arzuamouronte D, Herreros-Fernandez V. Health-related quality of life of subjects with known and unknown hypertension: results from the population-based Hortega study. *J Hypertens*, Jul 2003; 21(7): 1283-9.
- Menéndez E, Delgado E, Fernández-Vega F, Prieto MA, Bordiú E, Calle A, Carmena R, Castaño L, Catalá M, Franch J, Gaztambide S, Gírbés J, Goday A, Gomis R, López-Alba A, Martínez-Larrad MT, Mora-Peces I, Ortega E, Rojo-Martínez G, Serrano-Ríos MI, Urrutia I, Valdés S, Vázquez JA, Vendrell J, Soriguer F. Prevalence, Diagnosis, Treatment, and Control of Hypertension in Spain. Results of the Di@bet.es Study. *Rev Esp Cardiol (Engl Ed)*. 2016; 69(6):572-8.
- Minton JA1, Hattersley AT, Owen K, McCarthy MI, Walker M, Latif F, Barrett T, Frayling TM. Association studies of genetic variation in the WFS1 gene and type 2 diabetes in U.K. populations. *Diabetes*. 2002; 51(4):1287-90.
- Montague CT, Farooqi IS, Whitehead JP, Soos MA, Rau H, Wareham NJ, Sewter CP, Digby JE, Mohammed SN, Hurst JA, Cheetham CH, Earley AR, Barnett AH, Prins JB, O'Rahilly S. Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature* 1997; 387: 903–908.
- Moraes CT, Ciacci F, Bonilla E, Jansen C, Hirano M, Rao N, Lovelace RE, Rowland LP, Schon EA, DiMauro S. Two novel pathogenic mitochondrial ADN mutations affecting organelle number and protein synthesis. Is the tRNA (Leu(UUR)) gene an etiologic hot spot? *Journal of Clinical Investigation* 1993; 92: 2906-2915.
- Morcillo S, Martín-Núñez GM, Rojo-Martínez G, Almaraz MC, García-Escobar E, Mansego ML, De Marco G, Chaves FJ, Soriguer F. ELOVL6 genetic variation is related to insulin sensitivity: a new candidate gene in energy metabolism. *PLoS One*. 2011; 6(6):e21198.
- Morgan CL, Currie CJ, Peters JR. Relationship between diabetes and mortality: a population study using record linkage. *Diabetes Care* 2000; 23:1103-1107.
- Morris AP, Voight BF, Teslovich TM, et al.; DIAGRAM Consortium. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012;44:981–990.
- Morten KJ, Cooper JM, Brown GK, Lake BD, Pike D, Poulton J. A new point mutation associated with mitochondrial encephalomyopathy. *Human Molecular Genetics* 1993; 2(12): 2081-2087.

- Muller FL, Lustgarten MS, Jang Y, Richardson A, Van Remmen H. Trends in oxidative aging theories. *Free Radic Biol Med* 2007; 43:477–503.
- Munnich A, Rustin P. Clinical spectrum and diagnosis of mitochondrial disorders. *Am J Med Genet*. 2001 Spring; 106(1):4-17.
- Murea M, Ma L, Freedman BI. Genetic and environmental factors associated with type 2 diabetes and diabetic vascular complications. *Rev Diabet Stud*. 2012;9(1):6-22.
- Murphy R, Ellard S, Hattersley AT. Clinical implications of a molecular genetic classification of monogenic beta-cell diabetes. *Nat Clin Pract Endocrinol Metab*. 2008; 4(4):200-13.
- Naik RG, Brooks-Worrell BM, Palmer JP. Latent autoimmune diabetes in adults. *J Clin Endocrinol Metab*. 2009; 94(12):4635-44.
- Navarro A, Boveris A. The mitochondrial energy transduction system and the aging process. *Am J Physiol Cell Physiol* 2007; 292:C670–C686.
- Nestler JE, McClanahan MA. Diabetes and adrenal disease. *Baillieres Clin Endocrinol Metab*. 1992; 6(4):829-47.
- Nettleton JA, McKeown NM, Kanoni S, Lemaitre RN, Hivert MF, Ngwa J, van Rooij FJ, Sonestedt E, Wojczynski MK, Ye Z, Tanaka T, Garcia M, Anderson JS, Follis JL, Djousse L, Mukamal K, Papoutsakis C, Mozaffarian D, Zillikens MC, Bandinelli S, Bennett AJ, Borecki IB, Feitosa MF, Ferrucci L, Forouhi NG, Groves CJ, Hallmans G, Harris T, Hofman A, Houston DK, Hu FB, Johansson I, Kritchevsky SB, Langenberg C, Launer L, Liu Y, Loos RJ, Nalls M, Orholm-Melander M, Renstrom F, Rice K, Riserus U, Rolandsson O, Rotter JI, Saylor G, Sijbrands EJ, Sjogren P, Smith A, Steingrimsdottir L, Uitterlinden AG, Wareham NJ, Prokopenko I, Pankow JS, van Duijn CM, Florez JC, Witteman JC; MAGIC Investigators, Dupuis J, Dedoussis GV, Ordovas JM, Ingelsson E, Cupples L, Siscovick DS, Franks PW, Meigs JB. Interactions of dietary whole-grain intake with fasting glucose- and insulin-related genetic loci in individuals of European descent: a meta-analysis of 14 cohort studies. *Diabetes Care*. 2010; 33(12):2684-91.
- Olsson AH, Rönn T, Ladenvall C, Parikh H, Isomaa B, Groop L, Ling C. Two common genetic variants near nuclear-encoded OXPHOS genes are associated with insulin secretion in vivo. *Eur J Endocrinol*. 2011; 164(5): 765-71.
- Osbak KK, Colclough K, Saint-Martin C, Beer NL, Bellanné-Chantelot C, Ellard S, Gloy AL. Update on mutations in glucokinase (GCK), which cause maturity-onset diabetes of the young, permanent neonatal diabetes, and hyperinsulinemic hypoglycemia. *Hum Mutat* 2009; 30: 1512–1526.
- Palmer JP, Hirsch IB. What's in a name: latent autoimmune diabetes of adults, type 1.5, adult-onset, and type 1 diabetes. *Diabetes Care*. 2003; 26(2):536-8.
- Palmer ND, Goodarzi MO, Langefeld CD, Wang N, Guo X, Taylor KD, Fingerlin TE, Norris JM, Buchanan TA, Xiang AH, Haritunians T, Ziegler JT, Williams AH, Stefanovski D, Cui J, Mackay AW, Henkin LF, Bergman RN, Gao X, Gauderman J, Varma R, Hanis CL, Cox NJ, Highland HM, Below JE, Williams AL, Burt NP, Aguilar-Salinas CA, Huerta-Chagoya A, Gonzalez-Villalpando C, Orozco L, Haiman CA, Tsai MY, Johnson WC, Yao J, Rasmussen-Torvik L, Pankow J, Snively B, Jackson RD, Liu S, Nadler JL, Kandeel F, Chen YD, Bowden DW, Rich SS, Raffel LJ, Rotter JI, Watanabe RM, Wagenknecht LE. Genetic Variants Associated With Quantitative Glucose Homeostasis Traits Translate to Type 2 Diabetes in Mexican Americans: The GUARDIAN (Genetics Underlying Diabetes in Hispanics) Consortium. *Diabetes*. 2015; 64(5):1853-66.
- Palmieri L, Alberio S, Pisano I, Lodi T, Meznaric-Petrusa M, Zidar J, Santoro A, Scarcia P, Fontanesi F, Lamantea E, Ferrero I, Zeviani M. Complete loss-of-function of the heart/muscle-specific adenine nucleotide translocator is associated

with mitochondrial myopathy and cardiomyopathy. *Human Molecular Genetics* 2005; 14: 3079–3088.

- Parish R, Petersen KF. Mitochondrial dysfunction and type 2 diabetes. *Curr Diab Rep.* 2005 ;5(3):177-83.
- Patti ME, Butte AJ, Crunkhorn S, Cusi K, Berria R, Kashyap S, Miyazaki Y, Kohane I, Costello M, Saccone R, Landaker EJ, Goldfine AB, Mun E, DeFronzo R, Finlayson J, Kahn CR, Mandarino LJ. Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1. *Proc Natl Acad Sci USA* 2003; 100(14):8466-71.
- Patti ME, Corvera S. The role of mitochondria in the pathogenesis of type 2 diabetes. *Endocr Rev* 2010; 31: 364–95.
- Perry IJ. Healthy diet and lifestyle clustering and glucose intolerance. *Proc Nutr Soc.* 2002 Nov; 61(4):543-51.
- Perry JR, Voight BF, Yengo L, et al. MAGIC. DIAGRAM Consortium. GIANT Consortium Stratifying type 2 diabetes cases by BMI identifies genetic risk variants in LAMA1 and enrichment for risk variants in lean compared to obese cases. *PLoS Genet* 2012; 8:e1002741.
- Pigeyre M, Yazdi FT, Kaur Y, Meyre D. Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity. *Clin Sci (Lond)* 2016; 130(12):943-86.
- Pihlajamäki JI, Boes T, Kim EY, Dearie F, Kim BW, Schroeder J, Mun E, Nasser I, Park PJ, Bianco AC, Goldfine AB, Patti ME. Thyroid hormone-related regulation of gene expression in human fatty liver. *J Clin Endocrinol Metab.* 2009; 94(9):3521-9.
- Pollak FI, Vázquez T. Latent autoimmune diabetes in adults. *Rev Med Chil.* 2012 ; 140(11):1476-81.
- Poulsen P, Kyvik KO, Vaag A, Beck-Nielsen H. Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia.* 1999; 42(2):139-45.
- Poy MN, Eliasson L, Krutzfeldt J, Kuwajima S, Ma X, Macdonald PE, Pfeffer S, Tuschl T, Rajewsky N, Rorsman P, Stoffel M. A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* 2004; 432(7014):226-30.
- Pranavchand R, Reddy MM. Genomics era and complex disorders: Implications of GWAS with special reference to coronary artery disease, type 2 diabetes mellitus, and cancers. *J Postgrad Med.* 2016; 62(3):188-98.
- Prasad RB, Groop L. Genetics of type 2 diabetes-pitfalls and possibilities. *Genes (Basel)* 2015; 6(1):87-123.
- Pritchard JK, Cox NJ. The allelic architecture of human disease genes: common disease-common variant or not? *Hum. Mol. Genet.* 2002; 11:2417–2423.
- Procaccio V, Wallace, DC. Late-onset Leigh syndrome in a patient with mitochondrial complex I NDUFS8 mutations. *Neurology* 2004; 62: 1899–1901.
- Qi L, Cornelis MC, Zhang C, van Dam RM, Hu FB. Genetic predisposition, Western dietary pattern, and the risk of type 2 diabetes in men. *Am J Clin Nutr* 2009; 89:1453–1458.
- Quinlan A, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features *Bioinformatics.* 2010; 26(6): 841–842.
- R Development Core Team (2005) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Raciti GA, Longo M, Parrillo L, Ciccarelli M, Mirra P, Ungaro P, Formisano P, Miele C, Béguinot F. Understanding type 2 diabetes: from genetics to epigenetics. *Acta Diabetol.* 2015; 52(5):821-7.

- Rajagopalan S, Brook RD. Air pollution and type 2 diabetes: mechanistic insights. *Diabetes* 2012; 61:3037–45.
- Ramalingam L, Menikdiwela K, LeMieux M, Dufour JM, Kaur G, Kalupahana N, Moustaid-Moussa N. The renin-angiotensin system, oxidative stress and mitochondrial function in obesity and insulin resistance. *Biochim Biophys Acta*. 2016; pii: S0925-4439(16)30187-9.
- Rumpersaud E, Mitchell BD, Pollin TI, Fu M, Shen H, O'Connell JR, Ducharme JL, Hines S, Sack P, Naglieri R, Shuldiner AR, Snitker S. Physical activity and the association of common FTO gene variants with body mass index and obesity. *Arch Intern Med*. 2008; 168(16):1791-7.
- Reardon W, Ross RJ, Sweeney MG, Luxon LM, Pembrey ME, Harding AE, Trembath RC. Diabetes mellitus associated with a pathogenic point mutation in mitochondrial ADN. *Lancet* 1992; 340: 1376-1379.
- Report of a WHO/IDF consultation. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia. January 2006.
- Rice T, Pérusse L, Bouchard C, Rao DC. Familial aggregation of body mass index and subcutaneous fat measures in the longitudinal Quebec family study. *Genet Epidemiol* 1999; 16 : 316 – 334.
- Ritov VB, Menshikova EV, Azuma K, et al. Deficiency of electron transport chain in human skeletal muscle mitochondria in type 2 diabetes mellitus and obesity. *Am J Physiol Endocrinol Metab* 2010; 298: E49–E58.
- Rivière G, Lienhard D, Andrieu T, Vieau D, Frey BM, Frey FJ. Epigenetic regulation of somatic angiotensin-converting enzyme by ADN methylation and histone acetylation. *Epigenetics* 2011; 6:479–490.
- Rönn T, Poulsen P, Hansson O, Holmkvist J, Almgren P, Nilsson P, Tuomi T, Isomaa B, Groop L, Vaag A, Ling C. Age influences ADN methylation and gene expression of COX7A1 in human skeletal muscle. *Diabetologia* 2008; 51:1159–1168.
- Rönn T, Poulsen P, Tuomi T, Isomaa B, Groop L, Vaag A, Ling C. Genetic variation in ATP5O is associated with skeletal muscle ATP5O mRNA expression and glucose uptake in young twins. *PLoS ONE* 2009; 4:e4793.
- Rotig A, Munnich A. Genetic Features of Mitochondrial Respiratory Chain Disorders. *J Am Soc Nephrol* 2003; 14: 2995–3007.
- Rötig A, Munnich A. Genetic features of mitochondrial respiratory chain disorders. *J Am Soc Nephrol*. 2003; 14(12):2995-3007.
- Rottiers V, Näär AM. MicroRNAs in metabolism and metabolic disorders. *Nat Rev Mol Cell Biol* 2012; 13(4):239-50.
- Rozen S, Skaletsky H. Primer 3 on the WWW for general users and for biologist programmers. *Methods Mol Biol*. 2000; 132:365-386.
- Ruchat SM, Weisnagel SJ, Vohl MC, Rankinen T, Bouchard C, Pérusse L. Evidence for interaction between PPARG Pro12Ala and PPARGC1A Gly482Ser polymorphisms in determining type 2 diabetes intermediate phenotypes in overweight subjects. *Exp Clin Endocrinol Diabetes* 2009; 117:455-459.
- Rutter GA1, Chimienti F. SLC30A8 mutations in type 2 diabetes. *Diabetologia*. 2015; 58(1):31-6.
- Saada A, Shaag A, Mandel H, Nevo Y, Eriksson S, Elpeleg O. Mutant mitochondrial thymidine kinase in mitochondrial ADN depletion myopathy. *Nature Genetics* 2001; 29: 342–344.
- Saker PJ, Hattersley AT, Barrow B, Hammersley MS, Horton V, Gillmer MD, Turner RC. UKPDS 21: low prevalence of the mitochondrial transfer RNA gene (tRNA^{Leu}(UUR)) mutation at position 3243 bp in UK caucasian type 2 diabetic patients. *Diabetic Medicine* 1997; 14: 42-45.

- Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. *Nature Reviews Genetics* 2011; 12, 683-691.
- Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Althuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirm GW, Ma Q, Parikh H, Richardson D, Ricke D & Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007; 316: 1331-1336.
- Schapira AH. Mitochondrial diseases. *Lancet*. 2012; 379(9828):1825-34.
- Schiff M, Loublier S, Coulibaly A, Benit P, de Baulny HO, Rustin P. Mitochondria and diabetes mellitus: untangling a confl ictive relationship? *J Inherit Metab Dis* 2009; 32: 684-98.
- Schmidt O, Pfanner N, Meisinger C. Mitochondrial protein import: from proteomics to functional mechanisms. *Nat Rev Mol Cell Biol* 2010;11:655-667.
- Schwitzgebel VM. Many faces of monogenic diabetes. *J Diabetes Investig*. 2014;5(2):121-33.
- Simino J, Rao DC, Freedman BI. Novel findings and future directions on the genetics of hypertension. *Curr Opin Nephrol Hypertens*. 2012; 21(5):500-7.
- Sinnwell,J.P. and Schaid,D.J. (2005) , haplo.stats: statistical analysis of haplotypes with traits and covariates when linkage phase is ambiguous. R package version 1.2.2.
- Slomko H, Heo HJ, Einstein FH. Minireview: Epigenetics of obesity and diabetes in humans. *Endocrinology*. 2012; 153(3):1025-30.
- Smith PR, Bain SC, Good PA, Hattersley AT, Barnett AH, Gibson JM, Dodson PM. Pigmentary retinal dystrophy and the syndrome of maternally inherited diabetes and deafness caused by the mitochondrial ADN 3243 tRNA(Leu) A to G mutation. *Ophthalmology* 1999; 106: 1101-1108.
- Snogdal LS1, Wod M, Grarup N, Vestmar M, Sparsø T, Jørgensen T, Lauritzen T, Beck-Nielsen H, Henriksen JE, Pedersen O, Hansen T, Højlund K. Common variation in oxidative phosphorylation genes is not a major cause of insulin resistance or type 2 diabetes. *Diabetologia*. 2012; 55(2):340-8.
- Solé X, Guinó E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. *Bioinformatics*. 2006 Aug 1; 22(15):1928-9.
- Sonestedt E, Lyssenko V, Ericson U, Gullberg B, Wirfält E, Groop L, Orho-Melander M. Genetic variation in the glucose-dependent insulinotropic polypeptide receptor modifies the association between carbohydrate and fat intake and risk of type 2 diabetes in the Malmo Diet and Cancer cohort. *J Clin Endocrinol Metab* 2012; 97(5):E810-8.
- Song Y, Niu T, Manson JE, Kwiatkowski DJ, Liu S. Are variants in the CAPN10 gene related to risk of type 2 diabetes? A quantitative assessment of population and family-based association studies. *Am J Hum Genet* 2004; 74: 208-222.
- Soriguer F1, Goday A, Bosch-Comas A, Bordiú E, Calle-Pascual A, Carmena R, Casamitjana R, Castaño L, Castell C, Catalá M, Delgado E, Franch J, Gaztambide S, Girbés J, Gomis R, Gutiérrez G, López-Alba A, Martínez-Larrad MT, Menéndez E, Mora-Peces I, Ortega E, Pascual-Manich G, Rojo-Martínez G, Serrano-Rios M,

- Valdés S, Vázquez JA, Vendrell J. Prevalence of diabetes mellitus and impaired glucose regulation in Spain: the Di@bet.es Study. *Diabetologia*. 2012; 55(1):88-93.
- Soriguer-Escofet F, Esteva I, Rojo-Martinez G, Ruiz de Adana S, Catalá M, Merelo MJ, Aguilar M, Tinahones F, García-Almeida JM, Gómez-Zumaquero JM, Cuesta-Muñoz AL, Ortego J, Freire JM; Diabets Group of the Andalusian Society of Endocrinology and Nutrition. Prevalence of latent autoimmune diabetes of adults (LADA) in Southern Spain. *Diabetes Res Clin Pract*. 2002 Jun; 56(3):213-20.
 - Sparks LM, Xie H, Koza RA, Mynatt R, Hulver MW, Bray GA, Smith SR. A high-fat diet coordinately downregulates genes required for mitochondrial oxidative phosphorylation in skeletal muscle. *Diabetes* 2005; 54:1926–1933.
 - Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Allen HL, Lindgren CM, Luan J, Magi R, Randall JC, Vedantam S, Winkler TW, Qi L, Workalemahu T, Heid IM, Steinthorsdottir V, Stringham HM, Weedon MN, Wheeler E, Wood AR, Ferreira T, Weyant RJ, Segre AV, Estrada K, Liang L, Nemesh J, Park JH, Gustafsson S, Kilpelainen TO, Yang J, Bouatia-Naji N, Esko T, Feitosa MF, Kutalik Z, Mangino M, Raychaudhuri S, Scherag A, Smith AV, Welch R, Zhao JH, Aben KK, Absher DM, Amin N, Dixon AL, Fisher E, Glazer NL, Goddard ME, Heard- Costa NL, Hoesel V, Hottenga JJ, Johansson A, Johnson T, Ketkar S, Lamina C, Li S, Moffatt MF, Myers RH, Narisu N, Perry JR, Peters MJ, Preuss M, Ripatti S, Rivadeneira F, Sandholt C, Scott LJ, Timpson NJ, Tyrer JP, van Wingerden S, Watanabe RM, White CC, Wiklund F, Barlassina C, Chasman DI, Cooper MN, Jansson JO, Lawrence RW, Pellikka N, Prokopenko I, Shi J, Thiering E, Alavere H, Alibrandi MT, Almgren P, Arnold AM, Aspelund T, Atwood LD, Balkau B, Balmforth AJ, Bennett AJ, Ben-Shlomo Y, Bergman RN, Bergmann S, Biebertmann H, Blakemore AI, Boes T, Bonnycastle LL, Bornstein SR, Brown MJ, Buchanan TA, Busonero F, Campbell H, Cappuccio FP, Cavalcanti-Proenca C, Chen YD, Chen CM, Chines PS, Clarke R, Coin L, Connell J, Day IN, Heijer M, Duan J, Ebrahim S, Elliott P, Elosua R, Eiriksdottir G, Erdos MR, Eriksson JG, Facheris MF, Felix SB, Fischer-Posovszky P, Folsom AR, Friedrich N, Freimer NB, Fu M, Gaget S, Gejman PV, Geus EJ, Gieger C, Gjesing AP, Goel A, Goyette P, Grallert H, Grassler J, Greenawalt DM, Groves CJ, Gudnason V, Guiducci C, Hartikainen AL, Hassanali N, Hall AS, Havulinna AS, Hayward C, Heath AC, Hengstenberg C, Hicks AA, Hinney A, Hofman A, Homuth G, Hui J, Igl W, Iribarren C, Isomaa B, Jacobs KB, Jarick I, Jewell E, John U, Jorgensen T, Jousilahti P, Jula A, Kaakinen M, Kajantie E, Kaplan LM, Kathiresan S, Kettunen J, Kinnunen L, Knowles JW, Kolcic I, Konig IR, Koskinen S, Kovacs P, Kuusisto J, Kraft P, Kvaloy K, Laitinen J, Lantieri O, Lanzani C, Launer LJ, Lecoeur C, Lehtimaki T, Lettre G, Liu J, Lokki ML, Lorentzon M, Luben RN, Ludwig B, Manunta P, Marek D, Marre M, Martin NG, McArdle WL, McCarthy A, McKnight B, Meitinger T, Melander O, Meyre D, Midthjell K, Montgomery GW, Morken MA, Morris AP, Mulic R, Ngwa JS, Nelis M, Neville MJ, Nyholt DR, O'Donnell CJ, O'Rahilly S, Ong KK, Oostra B, Pare G, Parker AN, Perola M, Pichler I, Pietilainen KH, Platou CG, Polasek O, Pouta A, Rafelt S, Raitakari O, Rayner NW, Ridderstrale M, Rief W, Ruokonen A, Robertson NR, Rzehak P, Salomaa V, Sanders AR, Sandhu MS, Sanna S, Saramies J, Savolainen MJ, Scherag S, Schipf S, Schreiber S, Schunkert H, Silander K, Sinisalo J, Siscovick DS, Smit JH, Soranzo N, Sovio U, Stephens J, Surakka I, Swift AJ, Tammesoo ML, Tardif JC, Teder-Laving M, Teslovich TM, Thompson JR, Thomson B, Tonjes A, Tuomi T, van Meurs JB, van Ommen GJ, Vatin V, Viikari J, Visvikis-Siest S, Vitart V, Vogel CI, Voight BF, Waite LL, Wallaschofski H, Walters GB, Widén E, Wiegand S, Wild SH, Willemsen G, Witte DR, Wittteman JC, Xu J, Zhang Q, Zgaga L, Ziegler A, Zitting P, Beilby JP, Farooqi IS, Hebebrand J, Huikuri HV, James AL,

Kahonen M, Levinson DF, Macciardi F, Nieminen MS, Ohlsson C, Palmer LJ, Ridker PM, Stumvoll M, Beckmann JS, Boeing H, Boerwinkle E, Boomsma DI, Caulfield MJ, Chanock SJ, Collins FS, Cupples LA, Smith GD, Erdmann J, Froguel P, Gronberg H, Gyllenstein U, Hall P, Hansen T, Harris TB, Hattersley AT, Hayes RB, Heinrich J, Hu FB, Hveem K, Illig T, Jarvelin MR, Kaprio J, Karpe F, Khaw KT, Kiemeny LA, Krude H, Laakso M, Lawlor DA, Metspalu A, Munroe PB, Ouwehand WH, Pedersen O, Penninx BW, Peters A, Pramstaller PP, Quertermous T, Reinehr T, Rissanen A, Rudan I, Samani NJ, Schwarz PE, Shuldiner AR, Spector TD, Tuomilehto J, Uda M, Uitterlinden A, Valle TT, Wabitsch M, Waeber G, Wareham NJ, Watkins H, Wilson JF, Wright AF, Zillikens MC, Chatterjee N, McCarroll SA, Purcell S, Schadt EE, Visscher PM, Assimes TL, Borecki IB, Deloukas P, Fox CS, Groop LC, Haritunians T, Hunter DJ, Kaplan RC, Mohlke KL, O'Connell JR, Peltonen L, Schlessinger D, Strachan DP, van Duijn CM, Wichmann HE, Frayling TM, Thorsteinsdottir U, Abecasis GR, Barroso I, Boehnke M, Stefansson K, North KE, McCarthy MI, Hirschhorn JN, Ingelsson E, Loos RJ: Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. In *Nat Genet United States*, 2010, p. 937-948.

- Steinhorsdottir V, Thorleifsson G, Sulem P, Helgason H, Grarup N, Sigurdsson A, Helgadottir HT, Johannsdottir H, Magnusson OT, Gudjonsson SA, Justesen JM, Harder MN, Jørgensen ME, Christensen C, Brandslund I, Sandbæk A, Lauritzen T, Vestergaard H, Linneberg A, Jørgensen T, Hansen T, Daneshpour MS, Fallah MS, Hreidarsson AB, Sigurdsson G, Azizi F, Benediktsson R, Masson G, Helgason A, Kong A, Gudbjartsson DF, Pedersen O, Thorsteinsdottir U, Stefansson K. Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. *Nat Genet*. 2014;46(3):294-8.
- Stevens GA, Singh GM, Lu Y, Danaei G, Lin JK, Finucane MM, Bahalim AN, McIntire RK, Gutierrez HR, Cowan M, Paciorek CJ, Farzadfar F, Riley L, Ezzati M; Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Body Mass Index). National, regional, and global trends in adult overweight and obesity prevalences. *Popul Health Metr*. 2012 Nov 20;10(1):22.
- Stumvoll M, Goldstein BJ, van Haefen TW. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet*. 2005 Apr 9-15; 365(9467):1333-46.
- Stunkard AJ, Foch TT, Hrubec Z. A twin study of human obesity. *JAMA*. 1986; 256(1):51-4.
- Stunkard AJ, Sørensen TI, Hanis C, Teasdale TW, Chakraborty R, Schull WJ, Schulsinger F. An adoption study of human obesity. *N Engl J Med* 1986;314 : 193 – 198.
- Suzuki Y, Tsukuda K, Atsumi Y, Goto Y, Hosokawa K, Asahina T, Nonaka I, Matsuoka K, Oka Y. Clinical picture of a case of diabetes mellitus with mitochondrial tRNA mutation at position 3271. *Diabetes Care* 1996; 19: 1304-1305.
- Tallapragada DS, Bhaskar S, Chandak GR. New insights from monogenic diabetes for "common" type 2 diabetes. *Front Genet*. 2015; 6:251.
- Tange O. GNU Parallel - The Command-Line Power Tool. ;login: THE USENIX MAGAZINE 2011:42-47.
- Taurino F, Stanca E, Siculella L, Trentadue R, Papa S, Zanotti F, Gnoni A. Mitochondrial proteome analysis reveals depression of the Ndufs3 subunit and activity of complex I in diabetic rat brain. *J Proteomics*. 2012; 75(8):2331-41.
- Temelkova-Kurktschiev T, Stefanov T. Lifestyle and genetics in obesity and type 2 diabetes. *Exp Clin Endocrinol Diabetes*. 2012; 120(1):1-6.
- TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute, Crosby J, Peloso GM, Auer PL, Crosslin DR, Stitzel

- NO, Lange LA, Lu Y, Tang ZZ, Zhang H, Hindy G, Masca N, Stirrups K, Kanoni S, Do R, Jun G, Hu Y, Kang HM, Xue C, Goel A, Farrall M, Duga S, Merlini PA, Asselta R, Girelli D, Olivieri O, Martinelli N, Yin W, Reilly D, Speliotes E, Fox CS, Psaty BM, Holmen OL, Nikpay M, Farlow DN, Assimes TL, Franceschini N, Robinson J, North KE, Martin LW, DePristo M, Gupta N, Escher SA, Jansson JH, Van Zuydam N, Palmer CN, Wareham N, Koch W, Meitinger T, Peters A, Lieb W, Erbel R, König IR, Kruppa J, Degenhardt F, Gottesman O, Bottinger EP, O'Donnell CJ, Psaty BM, Ballantyne CM, Abecasis G, Ordovas JM, Melander O, Watkins H, Orho-Melander M, Ardissono D, Loos RJ, McPherson R, Willer CJ, Erdmann J, Hall AS, Samani NJ, Deloukas P, Schunkert H, Wilson JG, Kooperberg C, Rich SS, Tracy RP, Lin DY, Altshuler D, Gabriel S, Nickerson DA, Jarvik GP, Cupples LA, Reiner AP, Boerwinkle E, Kathiresan S. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N Engl J Med.* 2014 3;371(1):22-31.
- Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, Helgadóttir A, Styrkarsdóttir U, Gretarsdóttir S, Thorlacius S, Jonsdóttir I, Jonsdóttir T, Olafsdóttir EJ, Olafsdóttir GH, Jonsson T, Jonsson F, Borch-Johnsen K, Hansen T, Andersen G, Jorgensen T, Lauritzen T, Aben KK, Verbeek AL, Roeleveld N, Kampman E, Yanek LR, Becker LC, Tryggvadóttir L, Rafnar T, Becker DM, Gulcher J, Kiemeny LA, Pedersen O, Kong A, Thorsteinsdóttir U, Stefansson K. Genomewide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat Genet.* 2009; 41(1):18-24.
 - Tillil H, Köbberling J. Age-corrected empirical genetic risk estimates for first-degree relatives of IDDM patients. *Diabetes* 1987; 36: 93-99.
 - Tong Y, Lin Y, Zhang Y, Yang J, Zhang Y, Liu H, Zhang B. Association between TCF7L2 gene polymorphisms and susceptibility to type 2 diabetes mellitus: a large Human Genome Epidemiology (HuGE) review and metaanalysis. *BMC Med Genet* 2009; 10: 15.
 - Tzvetkov MV, Vormfelde SV, Balen D, Meineke I, Schmidt T, Sehr D, Sabolić I, Koepsell H, Brockmöller J. The effects of genetic polymorphisms in the organic cation transporters OCT1, OCT2, and OCT3 on the renal clearance of metformin. *Clin Pharmacol Ther.* 2009;86(3):299-306.
 - Valdés S, García-Torres F, Maldonado-Araque C, Goday A, Calle-Pascual A, Soriguer F, Castaño L, Catalá M, Gomis R, Rojo-Martínez G; Di@bet.es study group. Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Di@bet.es study. *Rev Esp Cardiol (Engl Ed).* 2014; 67(6):442-8.
 - Valdés S, Rojo-Martínez G, Soriguer F. Evolución de la prevalencia de la diabetes tipo 2 en población adulta española. *Med Clin (Barc).* 2007; 129(9):352-5.
 - van den Ouweland JM, Lemkes HH, Ruitenbeek W, Sandkuijl LA, de Vijlder MF, Struyvenberg PA, van de Kamp JJ, Maassen JA. Mutation in mitochondrial tRNA^{Leu}(uur) gene in a large pedigree with maternally transmitted type II diabetes mellitus and deafness. *Nature Genetics* 1992; 1: 368-371.
 - van Leeuwen N1, Nijpels G, Becker ML, Deshmukh H, Zhou K, Stricker BH, Uitterlinden AG, Hofman A, van 't Riet E, Palmer CN, Guigas B, Slagboom PE, Durrington P, Calle RA, Neil A, Hitman G, Livingstone SJ, Colhoun H, Holman RR, McCarthy MI, Dekker JM, 't Hart LM, Pearson ER. A gene variant near ATM is significantly associated with metformin treatment response in type 2 diabetes: areplication and meta-analysis of five cohorts. *Diabetologia.* 2012 ;55(7):1971-7.
 - Vandecasteele G, Szabadkai G, Rizzuto R. Mitochondrial calcium homeostasis: mechanisms and molecules. *IUBMB Life* 2001; 52: 213–219.

- Vaxillaire M, Bonnefond A, Froguel P. The lessons of early-onset monogenic diabetes for the understanding of diabetes pathogenesis. *Best Pract Res Clin Endocrinol Metab.* 2012; 26(2):171-87.
- Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, McCulloch LJ, Ferreira T, Grallert H, Amin N, Wu G, Willer CJ, Raychaudhuri S, McCarroll SA, Langenberg C, Hofmann OM, Dupuis J, Qi L, Segrè AV, van Hoek M, Navarro P, Ardlie K, Balkau B, Benediktsson R, Bennett AJ, Blagieva R, Boerwinkle E, Bonnycastle LL, Bengtsson Boström K, Bravenboer B, Bumpstead S, Burt NP, Charpentier G, Chines MR, Cornelis M, Couper DJ, Crawford G, Doney AS, Elliott KS, Elliott AL, Erdos PS, Fox CS, Franklin CS, Ganser M, Gieger C, Grarup N, Green T, Griffin S, Groves CJ, Guiducci C, Hadjadj S, Hassanali N, Herder C, Isomaa B, Jackson AU, Johnson PR, Jørgensen T, Kao WH, Klopp N, Kong A, Kraft P, Kuusisto J, Lauritzen T, Li M, Lieveise A, Lindgren CM, Lyssenko V, Marre M, Meitinger T, Midthjell K, Morken MA, Narisu N, Nilsson P, Owen KR, Payne F, Perry JR, Petersen AK, Platou C, Proença C, Prokopenko I, Rathmann W, Rayner NW, Robertson NR, Rocheleau G, Roden M, Sampson MJ, Saxena R, Shields BM, Shrader P, Sigurdsson G, Sparsø T, Strassburger K, Stringham HM, Sun Q, Swift AJ, Thorand B, Tichet J, Tuomi T, van Dam RM, van Haften TW, van Herpt T, van Vliet-Ostapchouk JV, Walters GB, Weedon MN, Wijmenga C, Witteman J, Bergman RN, Cauchi S, Collins FS, Gloyn AL, Gyllenstein U, Hansen T, Hide WA, Hitman GA, Hofman A, Hunter DJ, Hveem K, Laakso M, Mohlke KL, Morris AD, Palmer CN, Pramstaller PP, Rudan I, Sijbrands E, Stein LD, Tuomilehto J, Uitterlinden A, Walker M, Wareham NJ, Watanabe RM, Abecasis GR, Boehm BO, Campbell H, Daly MJ, Hattersley AT, Hu FB, Meigs JB, Pankow JS, Pedersen O, Wichmann HE, Barroso I, Florez JC, Frayling TM, Groop L, Sladek R, Thorsteinsdottir U, Wilson JF, Illig T, Froguel P, van Duijn CM, Stefansson K, Altshuler D, Boehnke M, McCarthy MI; MAGIC investigators; GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet.* 2011; 43(4): 388.
- Wallace DC, Fan W. Energetics, epigenetics, mitochondrial genetics. *Mitochondrion.* 2010; 10(1):12-31.
- Wallace, D.C. Why do we have a maternally inherited mitochondrial ADN? Insights from evolutionary medicine. *Annual Review of Biochemistry* 2007; 76, 781–821.
- Warnes, G. and Leisch, F. (2005) Genetics: Population Genetics. R package version 1.2.0.
- Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, Berndt SI, Elliott AL, Jackson AU, Lamina C, Lettre G, Lim N, Lyon HN, McCarroll SA, Papadakis K, Qi L, Randall JC, Roccascella RM, Sanna S, Scheet P, Weedon MN, Wheeler E, Zhao JH, Jacobs LC, Prokopenko I, Soranzo N, Tanaka T, Timpson NJ, Almgren P, Bennett A, Bergman RN, Bingham SA, Bonnycastle LL, Brown M, Burt NP, Chines P, Coin L, Collins FS, Connell JM, Cooper C, Smith GD, Dennison EM, Deodhar P, Elliott P, Erdos MR, Estrada K, Evans DM, Gianniny L, Gieger C, Gillson CJ, Guiducci C, Hackett R, Hadley D, Hall AS, Havulinna AS, Hebebrand J, Hofman A, Isomaa B, Jacobs KB, Johnson T, Jousilahti P, Jovanovic Z, Khaw KT, Kraft P, Kuokkanen M, Kuusisto J, Laitinen J, Lakatta EG, Luan J, Luben RN, Mangino M, McArdle WL, Meitinger T, Mulas A, Munroe PB, Narisu N, Ness AR, Northstone K, O'Rahilly S, Purmann C, Rees MG, Ridderstråle M, Ring SM, Rivadeneira F, Ruokonen A, Sandhu MS, Saramies J, Scott LJ, Scuteri A, Silander K, Sims MA, Song K, Stephens J, Stevens S, Stringham HM, Tung YC, Valle TT, Van Duijn CM, Vimalaswaran KS, Vollenweider P, Waeber G, Wallace C,

Watanabe RM, Waterworth DM, Watkins N; Wellcome Trust Case Control Consortium, Witterman JC, Zeggini E, Zhai G, Zillikens MC, Altshuler D, Caulfield MJ, Chanock SJ, Farooqi IS, Ferrucci L, Guralnik JM, Hattersley AT, Hu FB, Jarvelin MR, Laakso M, Mooser V, Ong KK, Ouwehand WH, Salomaa V, Samani NJ, Spector TD, Tuomi T, Tuomilehto J, Uda M, Uitterlinden AG, Wareham NJ, Deloukas P, Frayling TM, Groop LC, Hayes RB, Hunter DJ, Mohlke KL, Peltonen L, Schlessinger D, Strachan DP, Wichmann HE, McCarthy MI, Boehnke M, Barroso I, Abecasis GR, Hirschhorn JN. Genetic Investigation of anthropometric Traits Consortium. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet.* 2009; 41(1): 25-34.

- Wilson JF. In the clinic. Type 2 diabetes. *Ann Intern Med.* 2007 Jan 2; 146(1):ITC1-15; quiz ITC16.
- Wren JD, Garner HR. Data-mining analysis suggests an epigenetic pathogenesis for type 2 diabetes. *J Biomed Biotechnol* 2005; 2005(2):104-12.
- Wu C, Xu G, Tsai SA, Freed WJ, Lee CT. Transcriptional profiles of type 2 diabetes in human skeletal muscle reveal insulin resistance, metabolic defects, apoptosis, and molecular signatures of immune activation in response to infections. *Biochem Biophys Res Commun.* 2017; 482(2):282-288.
- Wu HH, Liu NJ, Yang Z, Tao XM, Du YP, Wang XC, Lu B, Zhang ZY, Hu RM, Wen J. IGF2BP2 and obesity interaction analysis for type 2 diabetes mellitus in Chinese Han population. *Eur J Med Res.* 2014; 19:40.
- Xia Q, Grant SF. The genetics of human obesity. *Ann N Y Acad Sci.* 2013; 1281:178-90.
- Xu X, Jiang SY, Wang TY, Bai Y, Zhong M, Wang A, et al. Inflammatory response to fine particulate air pollution exposure: neutrophil versus monocyte. *PLoS One* 2013; 8:e71414.
- Yeo GS, Connie Hung CC, Rochford J, Keogh J, Gray J, Sivaramakrishnan S, O'Rahilly S, Farooqi IS. A de novo mutation affecting human TrkB associated with severe obesity and developmental delay. *Nat. Neurosci* 2004, 7: 1187–1189.
- Youle RJ, Karbowski M. Mitochondrial fission in apoptosis. *Nature Reviews Molecular Cell Biology* 2005; 6: 657–663.
- Zeggini E1, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Boström KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jørgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marville AF, Meisinger C, Midtthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjögren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ; Wellcome Trust Case Control Consortium, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nature genetics.* 2008; 40(5):638-45.
- Zhao Q, Kelly TN, Li C, He J. Progress and future aspects in genetics of human hypertension. *Curr Hypertens Rep.* 2013; 15(6):676-86.

- Zhu X, Peng X, Guan MX, Yan Q. Pathogenic mutations of nuclear genes associated with mitochondrial disorders. *Acta Biochim Biophys Sin (Shanghai)* 2009; 41(3):179-87.
- Zhu Z, Yao J, Johns T, Fu K, De Bie I, Macmillan C, Cuthbert AP, Newbold RF, Wang J, Chevrette M, Brown GK, Brown RM, Shoubridge EA. SURF1, encoding a factor involved in the biogenesis of cytochrome c oxidase, is mutated in Leigh syndrome. *Nature Genetics* 1998; 20: 337–343.
- Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. *Nature*. 2001; 414(6865):782-7.
- Zuchner S, Mersiyanova IV, Muglia M, Bissar-Tadmouri N, Rochelle J, Dadali EL, Zappia M, Nelis E, Patitucci A, Senderek J, Parman Y, Evgrafov O, Jonghe PD, Takahashi Y, Tsuji S, Pericak-Vance MA, Quattrone A, Battaloglu E, Polyakov AV, Timmerman V, Schroder JM, Vance JM. Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot-Marie-Tooth neuropathy type 2A. *Nature Genetics* 2004; 36: 449–451.

9. Anexo

Anexo I. Información de los 96 polimorfismos estudiados en la fase 1 de este trabajo.

| SNP | Locus | Ch r | chrcoord | Consecuencia | Nombre HGVS | MAF Ncbi / real |
|-----------------|----------|----------|---------------|---------------------------------------|--------------------------------------|-------------------------------|
| rs33946522 | BCS1L | 2 | 21923595 6 | SYNONYMOUS_CODING | NM_001079866.1:c.996C>T (p.N332N) | 0.076/ND |
| rs2243911 | | | 21923759 2 | INTRONIC | NM_022453.2:c.802-90G>A | 0.496/0.461 |
| rs12190858 0 | | | 21952585 8 | MISSENSE | NM_004328.4:c.148A>G (p.Thr50Ala) | ND/0 (OMIN60364 7.0011) |
| rs11557185 | COX4I1 | 16 | 84396082 | NON_SYNONYMOUS_CO DING | NM_001861.2:c.112T>C (p.Y38H) | ND/0 |
| rs11648858 | | | 84389793 | INTRONIC | NM_006067.3:c.231+379G>C | 0.145/0 |
| rs6495131 | COX5A | 15 | 73017555 | UPSTREAM | NT_010194.16:g.46020820A> G | 0.273/0 |
| rs11553378 | | | 73008597 | NON_SYNONYMOUS_CO DING | NM_004255.2:c.130G>C (p.G44R) | ND/0 |
| rs4131311 | | | 73006160 | NON_SYNONYMOUS_CO DING-SPLICE_SITE | NM_004255.2:c.339G>T (p.K113N) | 0.018/0 |
| rs11553377 | | | 73008579 | STOP_GAINED | NM_004255.2:c.148G>T (p.E50*) | ND/0 |
| rs1133322 | | | 72999410 | DOWNSTREAM | NT_010194.16:g.46002675A> G | 0.318/0.484 |
| rs1470625 | | | COX5B | 2 | 97631923 | DOWNSTREAM |
| rs17022045 | 97628120 | UPSTREAM | | | NT_022171.14:g.331273C>G | 0.122/0.120 |
| rs17431357 | COX6A1 | 12 | 11936481 7 | UPSTREAM | NT_009775.16:g.11449943C> T | 0.030/0.062 |
| rs12310837 | | | 11936495 5 | UPSTREAM | NT_009775.16:g.11450081A> G | 0.141/0.212 |
| rs2076022 | | | 11936017 3 | UPSTREAM | NT_009775.17:g.11452320C> T | 0.488/0.353 |
| rs4775 | | | 11936264 1 | NON_SYNONYMOUS_CO DING | NT_009775.16 g.11447767C>T | ND/0 |
| rs4889657 | COX6A2 | 16 | 31344433 | INTRONIC (UPSTREAM) | NM_005353.2:c.3373- 404C>T | 0.086/0 |
| rs4806187 | COX6B1 | 19 | 40841601 | DOWNSTREAM | NT_011109.15:g.8417979A> G | 0.381/0.491 |
| rs1058511 | COX6B2 | 19 | 60551251 | 3PRIME_UTR | NM_032701.2:c.*622C>A | 0.4619/0.474 |
| rs1130569 | COX6C | 8 | 10096896 9 | SYNONYMOUS_CODING | NM_004374.2:c.168C>T(p.Y 56Y) | 0.267/0.151 |
| rs4626565 | | | 10095971 8 | INTRONIC | NM_004374.2:c.16-49A>G | 0.242/0.152 |
| rs17882491 | COX7A1 | 19 | 41334261 | NON_SYNONYMOUS_CO DING | NM_001864.2:c.130A>C (p.K44Q) | 0.001/0 |
| rs753420 | | | 41335535 | 5PRIME_UTR | NM_001864.2:c.-386A>C | 0.394/0.266 |
| rs7255180 | | | 41332159 | INTRONIC | NM_001003962.1:c.722- 161C>T | 0.029/0.061 |
| rs436898 | COX7A2 | 6 | 76002268 | DOWNSTREAM | NT_007299.12:g.13765720A> C | 0.142/0.073 |
| rs683943 | | | 76018528 | 5' UTR | NT_007299.12:g.13781980C> G | 0.037/0.045 |
| rs9360898 | | | 76010425 | 5PRIME_UTR | NT_007299.12:g.13773877T> G | 0.202/0.221 |
| rs9790574 | COX7B2 | 4 | 46429658 | DOWNSTREAM | NT_006238.10:g.6436867C> T | 0.224/0.173 |
| rs16902285 | COX7C | 5 | 85953326 | DOWNSTREAM | NT_006713.14:g.36511928T> G | 0.052/0.054 |
| rs13161296 | | | 85948806 | UPSTREAM | NT_006713.14:g.36507408C> T | 0.013/0.017 |

| | | | | | | |
|-------------|--------|----|-----------|-----------------------|--------------------------------------|--------------|
| rs2410718 | | | 85954228 | DOWNSTREAM | NT_006713.14:g.36512830G>A | 0.124/0.042 |
| rs2089095 | COX8C | 14 | 92886692 | INTRONIC | NM_020818.3:c.-351+17070C>G | 0.098/0.103 |
| rs1053517 | NDUFS1 | 2 | 206695260 | DOWNSTREAM | NT_005403.16:g.57196432A>G | 0.481/0.493 |
| rs3770989 | | | 206696996 | 3PRIME_UTR | NM_005006.5:c.*158T>C | 0.023/0.039 |
| rs11548670 | | | 206720728 | SYNONYMOUS_CODING | NM_005006.5:c.414T>C (p.D138D) | 0.054/0.042 |
| rs4656994 | NDUFS2 | 1 | 159446501 | INTRONIC | NM_004550.3:c.703-24G>A | 0.277/0.241 |
| rs12094497 | | | 159452927 | INTRONIC | NM_004106.1:c.49+1143G>A | 0.0675/0.065 |
| rs1136224 | | | 159450721 | 3PRIME_UTR | NM_004550.3:c.*114A>G | 0.166/0.136 |
| rs1136207 | | | 159450140 | SYNONYMOUS_CODING | NM_004550.3:c.1290C>T (p.A430A) | 0.137/0.170 |
| rs4656993 | | | 159442761 | INTRONIC | NM_004550.3:c.203-60A>G | 0.307/0.371 |
| rs11587213 | | | 159451499 | UPSTREAM | NT_004487.18:g.11675230A>G | 0.141/0.177 |
| rs6677859 | | | 159448873 | SYNONYMOUS_CODING | NM_004550.3:c.1095A>G (p.P365P) | ND/0 |
| rs9600 | NDUFS3 | 11 | 47562560 | NON_SYNONYMOUS_CODING | NM_004551.1:c.746C>A (p.P249Q) | ND/0 |
| rs3740654 | | | 47562561 | SYNONYMOUS_CODING | NM_004551.1:c.747G>A (p.P249P) | 0.002/0 |
| rs10742816 | | | 47563560 | UPSTREAM | NT_009237.17:g.46394225C>T | 0.272/0.087 |
| rs10838732 | | | 47564516 | UPSTREAM | NT_009237.17:g.46395181C>G | 0.351/0.329 |
| rs2280231 | | | 47557014 | INTRONIC | NM_016506.4:c.-88G>A | 0.193/0.189 |
| rs121908985 | NDUFS4 | 5 | 52942176 | STOP_GAINED | NM_002495.2:c.291delG (p.Trp97Terfs) | ND/0 |
| rs31304 | | | 52977840 | SYNONYMOUS_CODING | NM_002495.1:c.198A>C (p.G66G) | 0.049/0.028 |
| rs567 | | | 53014854 | 3PRIME_UTR | NM_002495.1:c.*46G>A | 0.351/0.488 |
| rs1044692 | | | 53014800 | NON_SYNONYMOUS_CODING | NM_002495.1:c.520A>C (p.T174P) | ND/0 |
| rs535277 | | | 52990271 | INTRONIC | NM_002495.1:c.424+60T>G | 0.303/0.167 |
| rs1984600 | NDUFS5 | 1 | 39264514 | UPSTREAM | NT_032977.8:g.9463845G>C | 0.395/0.181 |
| rs11205591 | | | 39273836 | DOWNSTREAM | NT_032977.8:g.9473167C>G | 0.252/0.267 |
| rs10058270 | NDUFS6 | 5 | 1868950 | INTRONIC | NM_004553.3:c.310-15T>C | 0.119/0.093 |
| rs11953620 | | | 1871630 | INTRONIC | NT_006576.15:g.1808630C>T | 0.396/0.318 |
| rs4975851 | | | 1870336 | INTRONIC | NT_006576.15:g.1807336C>T | 0.371/0.117 |
| rs35349407 | | | 1867479 | NON_SYNONYMOUS_CODING | NM_004553.3:c.213G>C (p.L71F) | 0.013/0 |
| rs972891 | | | 1853812 | UPSTREAM | NT_006576.15:g.1790812G>T | 0.247/0 |
| rs11666067 | NDUFS7 | 19 | 1340517 | NON_SYNONYMOUS_CODING | NM_024407.3:c.228+580C>A | 0.476/0.374 |
| rs1142530 | | | 1339538 | NON_SYNONYMOUS_CODING | NM_024407.3:c.68C>T (p.P23L) | 0.456/0.444 |
| rs11668291 | | | 1335246 | INTRONIC | NM_024407.3:c.16+305C>T | 0.010/.015 |
| rs809359 | | | 1342809 | INTRONIC | NM_024407.3:c.455+645G>A | 0.060/0.082 |

| | | | | | | |
|-------------|--------|-------------------------------|-----------|-----------------------------------|---------------------------------|-----------------------|
| rs2074896 | | | 1342059 | SPLICE_SITE-INTRONIC | NM_024407.3:c.408+10G>T | 0.122/0 |
| rs3848638 | | | 1334649 | UPSTREAM | NT_011255.14:g.1323649G>T | 0.381/0 |
| rs4147780 | NDUFS8 | 11 | 67561957 | UPSTREAM | NT_033903.7:g.13111176G>A | 0.491/0.490 |
| rs11556013 | | | 67556995 | NON_SYNONYMOUS_CODING | NM_002496.1:c.139G>T (p.D47Y) | ND/0 |
| rs4147779 | | | 67561383 | UPSTREAM | NT_033903.7:g.13110602G>A | 0.253/0.257 |
| rs10896288 | | | 67554281 | UPSTREAM | NT_033903.7:g.13103500C>G | 0.467/0.464 |
| rs1061519 | | | SDHA | 5 | 276647 | NON_SYNONYMOUS_CODING |
| rs1061520 | 276654 | NON_SYNONYMOUS_CODING | | | NM_004168.2:c.121A>G (p.K41E) | ND/0 |
| rs7710005 | 286572 | INTRONIC | | | NM_004168.2:c.896-20A>G | 0.215/0.015 |
| rs34635677 | 276646 | NON_SYNONYMOUS_CODING | | | NM_004168.2:c.113A>T (p.D38V) | 0.017/0.012 |
| rs6555055 | 279160 | SPLICE_SITE-SYNONYMOUS_CODING | | | NM_004168.2:c.619A>C (p.R207R) | 0.247/0.215 |
| rs6962 | 309509 | NON_SYNONYMOUS_CODING | | | NM_004168.2:c.1969G>A (p.V657I) | 0.175/0.161 |
| rs1022580 | SDHB | 1 | 17232263 | INTRONIC | NM_003000.2:c.201-36G>T | 0.043/0.033 |
| rs786201095 | | | 17355138 | MISSENSE | NM_003000.2:c.380T>A (I127N) | ND/0 |
| rs2647169 | | | 17243810 | INTRONIC | NM_003000.2:c.200+33G>A | 0.105/0.065 |
| rs11203280 | | | 17215370 | UPSTREAM NO ENSEML | NT_004610.18:g.167125G>T | 0.482/0.411 |
| rs74315367 | | | 17350520 | MISSENSE | NM_003000.2:c.590C>G (P197R) | ND/0 |
| rs34916635 | | | 17243885 | NON_SYNONYMOUS_CODING | NM_003000.2:c.158G>A (p.G53E) | 0.022/0 |
| rs35962811 | | | 17243873 | NON_SYNONYMOUS_CODING | NM_003000.2:c.170A>G | 0.001/0 |
| rs33927012 | | | 17226884 | NON_SYNONYMOUS_CODING | NM_003000.2:c.487T>C (p.S163P) | 0.008/0.007 |
| rs34599281 | | | 17243865 | NON_SYNONYMOUS_CODING | NM_003000.2:c.178A>G (p.T60A) | 7.10 ⁻⁵ /0 |
| rs9435739 | | | 17215511 | UPSTREAM NO ENSEML | NT_004610.18:g.167266G>A | 0.478/0.394 |
| CI034610 | | | | | 17349206 | MISSENSE |
| rs4600063 | SDHC | 1 | 159599815 | 3PRIME_UTR | NM_001035511.1:c.*861A>G | 0.142/0.062 |
| rs13374037 | | | 159601905 | INTRONIC | NM_001013625.2:c.348+35A>T | 0.142/0.067 |
| rs10891319 | SDHD | 11 | 111472815 | DOWNSTREAM | NT_033899.7:g.15530021A>G | 0.368/0.296 |
| rs11088 | UQCRB | 8 | 97313341 | INTRONIC | NT_008046.15 g.10462338A>T | ND/0 |
| rs10504961 | | | 97309305 | 3' UTR | NM_006294.3:c.*3154G>A | 0.376/0.470 |
| rs35895613 | | | 97314562 | NON_SYNONYMOUS_CODING-SPLICE_SITE | NM_006294.2:c.89T>C (p.L30P) | 0.041/0 |
| rs803224 | UQCRQ | 5 | 132234403 | UPSTREAM | NT_034772.5:g.34621517G>A | 0.190/0.102 |
| rs17624157 | | | 132231930 | UPSTREAM | NM_014402.4:c.*757G>A | 0.074/0.154 |
| rs39830 | UQCRQ | 5 | 132229317 | UPSTREAM | NT_034772.5:g.34616431G>A | 0.447/0.404 |
| rs30177 | | | 132230310 | 5PRIME_UTR | NM_014402.4:c.-14+32C>G | 0.405/0.251 |

Anexo II. Cebadores utilizados en la PCR1 de la fase 3 del trabajo, para cada variante secuenciada, y tamaños de los amplicones de la PCR1 y PCR2.

| SNP amplificado | Secuencias de los cebadores específicos para los SNPs secuenciados (incluyen los adaptadores) | Tamaño del producto de PCR1 | Tamaño del producto de PCR2 |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|-----------------------------|
| rs10157617 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGTCTCTGTAATTG ATGGGTTTGG_5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGGCAACAGAGC AAGACCTTATC_3' | 185 | 254 |
| rs1801318 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAAGATAAATTGGA ACGCAATCTGT_5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGAAGGCAATGG GTTTGTATGATAG_3' | 189 | 258 |
| rs4147729 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAGCTAAGCAGTTA CTGGTCTCATCA_5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGCAACACTGTTT TGAAAGGTACG_3' | 194 | 263 |
| rs6981 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGGAAAGTACACC CCTCCACCT_5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGGTTTGTCTGTA ACAACCTTTC_3' | 200 | 269 |
| rs1122472 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTCGCTAGCTAGTTC GTTCTGAAGC_5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGCAAGCCAAGA GGAAATCTGTTA_3' | 215 | 284 |
| rs77655991 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTATCATTTCCTCTG CCGTGTGTTTC_5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTGTTTGGAAAAT CATCTCTTGTG_3' | 218 | 287 |
| rs4147718 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGAAGATGCCATTT TTCACATATAA_5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAAGAGGCAGAT GTTGTCTCTG_3' | 225 | 294 |
| rs1984600 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAGTGATCTTCCTC CCTCAGCCTTC_5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTAAGGATGACC TCAGCGAGAC_3' | 229 | 298 |
| rs2746462 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCTGAGGCTCC AGGACTCAC_5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCGATGTTTCGACG GGACAC_3' | 233 | 302 |

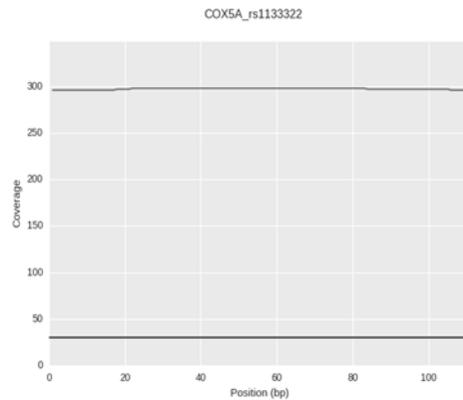
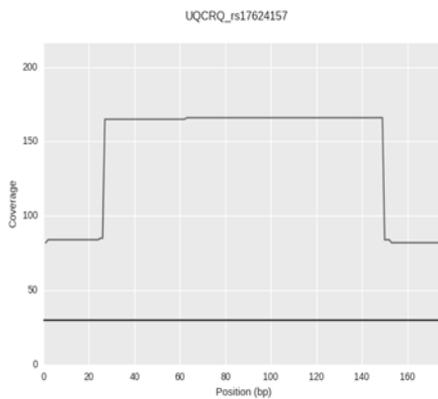
Anexo III. Cebadores utilizados en la PCR1 de la fase 4 del trabajo, para cada variante secuenciada, y tamaños de los amplicones de la PCR1 y PCR2.

| Fragmento amplificado | Secuencias de los cebadores específicos para los SNPs secuenciados (incluyen los adaptadores) | Tamaño del producto de PCR1 | Tamaño del producto de PCR2 |
|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|-----------------------------|
| rs10891319 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCCCCACCTTGTG GTGTCATTTC 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTTCATCACTCCA CTCTACAGCAG 3' | 142 | 210 |
| rs39830 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTATAAAACACAG CCAACTGTGAA 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAGTGTTTTGTTA ATGGAAATGCTG 3' | 147 | 215 |
| rs17624157 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCTGGGCCCTAAA TACTCTTCTT 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTTTAAGAGCATC ACTTTAGTGCAA 3 | 242 | 310 |
| rs803224 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTTCAGTCTGGTA ACCCATCTG 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGGACTGCCTGCA AACTAAAAGAG 3' | 259 | 327 |
| rs4600063 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGGAATCAGATAG AACTTAGCCAGA 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCATGCGTCACTTC TTTTCTTTTCT 3' | 154 | 222 |
| rs13374037 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGGTGGCATAACA ATAGACATTTT 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGTCTGAAATCT TAGGCAAGGTAT 3' | 172 | 240 |
| rs1136224 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCACTGGAAATG GCCTCTGT 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTACATGCACAGA AAGCCTGACA 3' | 159 | 227 |
| rs4656993 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTCTCTTCCTGGT TATAGTGGAGA 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGTTCAGAGCCAC ATGGTTATTCC 3' | 202 | 270 |
| rs7255180 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGACACTTATCTCA CAGAGTTGGATT 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAGCTGGCACTTC ACTGAGTATTT 3' | 166 | 234 |
| rs1133322 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCACATCCTCAAGA CTATCTCAGTG 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTAGGCTTAGGGG ACTGGTTGTC 3' | 177 | 245 |
| rs10742816 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAGTTCTCAGATC CACAGGGTTC 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAAAAAGAACCCC ACCATCTCAG 3' | 181 | 249 |
| rs17431357 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCAGCCAGCTGGTT TATCTTT 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATATGCCTTACTT TTGGCTTACCA 3' | 184 | 252 |
| rs683943 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAATATGGGGAGG GACAGATACAA 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAAGGGGTCAGG GAGGTTCTCTTA 3' | 187 | 255 |
| rs9360898 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAAAAGAAAAACT AGAGCGGCAGT 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACATTCGTCTCAT AAATCGTTTGC 3' | 221 | 289 |
| rs1470625 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAGCCCACTTTAGC TGCTTCATAC 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTAACTGAAGCCAC CTCTGGCTAAT 3' | 194 | 262 |
| rs1053517 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAGTATCTTCTG CCTGAACC 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCTTTTCTGAT ACAACAAAGGAT 3' | 198 | 266 |
| rs3770989 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTTACACACAAT ACATGTTTTTCA 5' GTCCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGTGATCCTTTAC AGGTTTATTTCTTG 3' | 212 | 280 |
| rs2410718 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGGCTGTAAGTGGC | 206 | 274 |

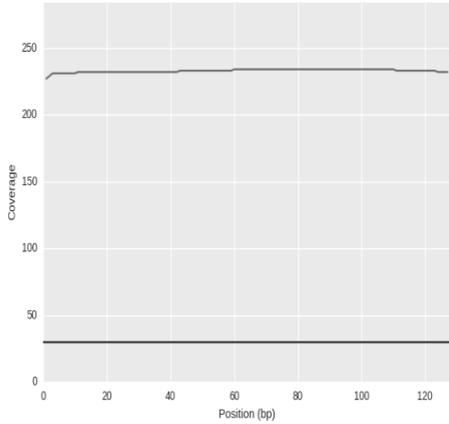
| | | | |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----|
| | TAAGTTTATTG 5' GTC TCG TGGGCTCGGAGATGTGTATAAGAGACAGCAAGCCTCCAGAA TCAGATGACA 3' | | |
| rs535277 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGATGCAGTTTCCT TTGCAGAAA 5' GTC TCG TGGGCTCGGAGATGTGTATAAGAGACAGGGATTAGATACAA TGAAACCAC TGG A 3' | 217 | 285 |
| rs4806187 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAGTGAGACTATGC GTGTGAACG 5' GTC TCG TGGGCTCGGAGATGTGTATAAGAGACAGTGAAGAACGAGA CTCCATCATAAAA 3' | 230 | 298 |
| rs11205591 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTTACTCTTTCACC CAGCCTGA 5' GTC TCG TGGGCTCGGAGATGTGTATAAGAGACAGCCAGCTCAAGTCT ATGTTTTATGA 3' | 239 | 307 |
| rs10504961 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTTAAAGTAAATA ACTAAATGAAATCACTACCA 5' GTC TCG TGGGCTCGGAGATGTGTATAAGAGACAGACATTCTGC ACT GAGTTATCATGG 3' | 246 | 314 |
| rs1022580 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCTTCGGAAGGTCA AAGTAGAGTCA 5' GTC TCG TGGGCTCGGAGATGTGTATAAGAGACAGCCTGAGAAGACCA AATGGATAAGC 3' | 254 | 322 |
| rs1801318 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAAGATAATTGGA ACGCAAAATCTGT_5' GTC TCG TGGGCTCGGAGATGTGTATAAGAGACAGAGAAGGCAATGG GTTTATGATAG_3' | 189 | 258 |
| rs77655991 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTATCATTTCCCTCG CCTGTTGTC_5' GTC TCG TGGGCTCGGAGATGTGTATAAGAGACAGGTGTTTGGAAAAAT CATCTTTTGTG_3' | 218 | 287 |
| rs2746462 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCTCTGAGGCCTCC AGGACTCAC_5' GTC TCG TGGGCTCGGAGATGTGTATAAGAGACAGGCGATGTTTCGACG GGACAC_3' | 233 | 302 |

Anexo IV. Coberturas de los 26 fragmentos secuenciados en la fase 4 del trabajo.

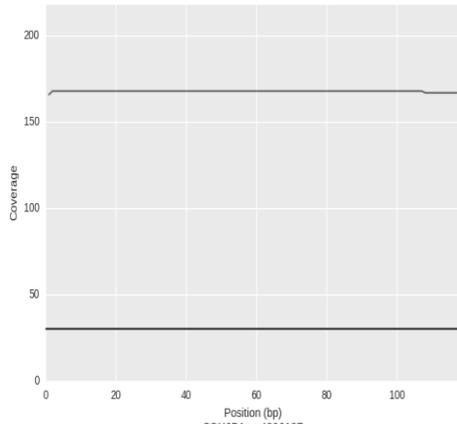
rs274: rs2746462; rs776: rs77655991; rs7718: rs1801318



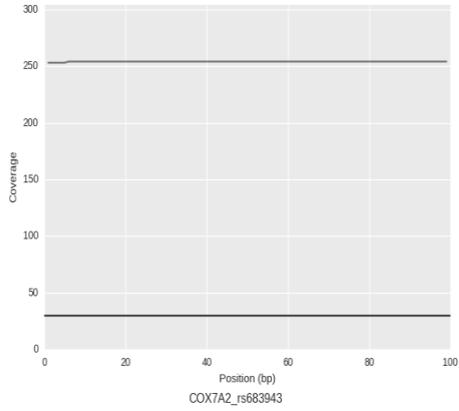
COX5B_rs1470625



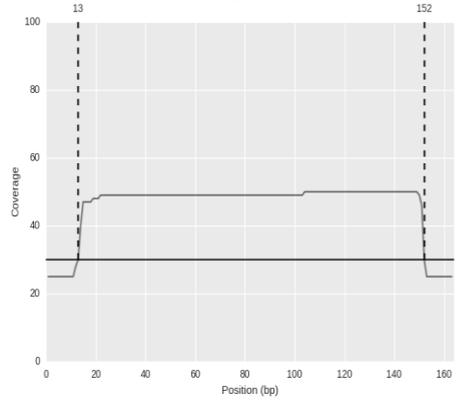
COX6A1_rs17431357



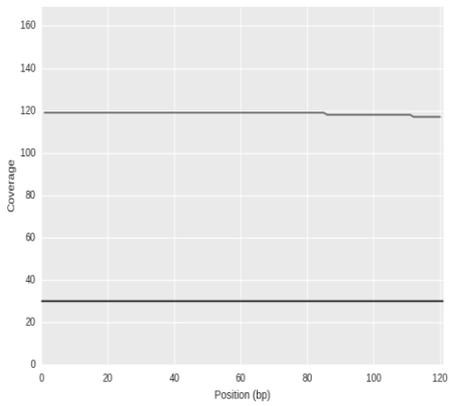
COX7A1_rs7255180



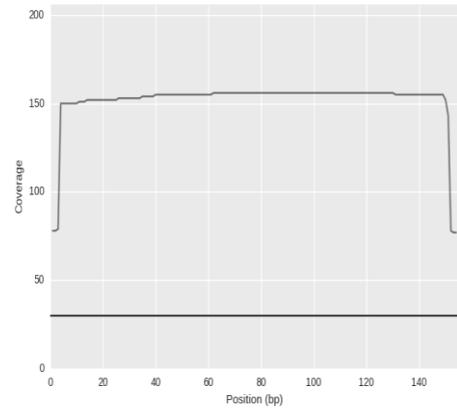
COX6B1_rs4806187



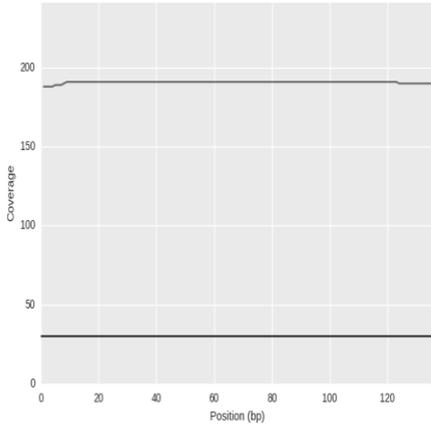
COX7A2_rs683943



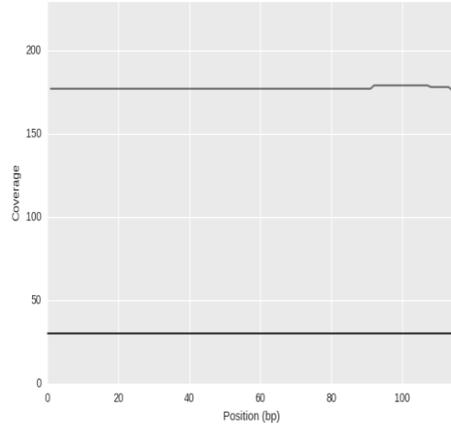
COX7A2_rs9360898



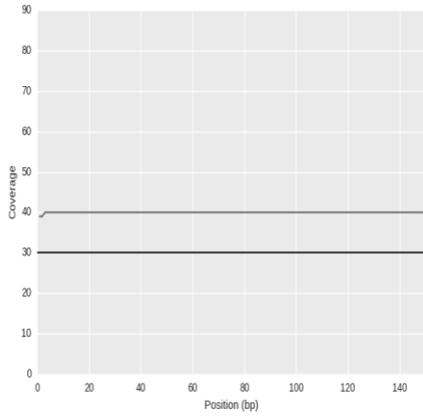
NDUFS2_rs4656993



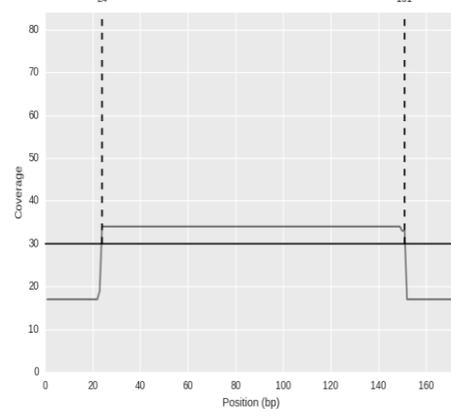
NDUFS3_rs10742816



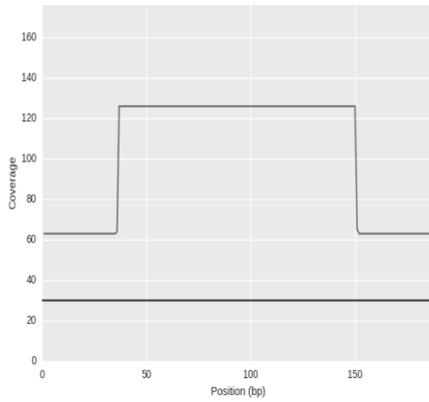
NDUFS4_rs535277



NDUFS5_rs11205591



SDHB_rs1022580



SDHB-rs274

