# PROGRAMA DE DESARROLLO DE CIENCIAS BÁSICAS PEDECIBA – ÁREA BIOLOGÍA DOCTORADO EN CIENCIAS BIOLÓGICAS

# Metilación diferencial del ADN de leucocitos en pacientes con cáncer de mama y melanoma:

# nuevo abordaje en la búsqueda de biomarcadores de riesgo

Mag. Mónica Cappetta Sapriza

**ORIENTADORES:** 

Dra. María Berdasco

Dr. Bernardo Bertoni

**MONTEVIDEO, 13 DE MAYO DE 2014** 

A Mica, Guille y Walter

Gracias por permitirme disfrutar todos los días de su amor, su paciencia y la alegría que aportan a mi vida.

#### AGRADECIMIENTOS

A *Bernardo Bertoni*, por confiar en mí y darme el espacio y apoyo para iniciarme en el mundo de la epigenética. Corrimos un riesgo enorme que por suerte salió bien. Gracias por dejarme crecer y alentarme todos los días a cumplir mis metas y no achicarme.

A *María Berdasco*, que a pesar de la distancia siempre estuviste presente para responder todas mis dudas, apoyarme y enseñarme un montón a través de cientos de mails, horas de chateo y mis instancias en el Idibell. Gracias por la oportunidad que me diste.

A *los "isleños", Silvana, Valentina, Lucía, Soledad, Jorge, Tatiana y Jimena*, con quienes no solo trabajo todos los días, sino compartimos alegrías, tristezas, dudas, salidas y muchas risas. Gracias por hacer del laboratorio un lugar tan ameno para trabajar y sentir el apoyo de todos ustedes todos los días. Son mucho más que compañeros de trabajo.

A todos *mis compañeros del Departamento de Genética* por su apoyo constante y por compartir conmigo tantos momentos gratos dentro y fuera del trabajo. En especial a *Lorena Da Silveira*, quien trabajó mano a mano conmigo, siendo una enorme ayuda en la puesta a punto de la técnica de AIMS.

A *Leda Roche, Pepe Tort y María Mirta Rodríguez*, integrantes del tribunal de mi concurso de grado 2, por haber elegido para la prueba el primer artículo científico sobre epigenética que leí con detenimiento. Sin saberlo desencadenaron mi interés por la epigenética.

A todos los *integrantes del proyecto Komen*, sin ustedes todo este trabajo no se hubiera podido llevar a cabo. Gracias por permitirme utilizar los datos y las muestras para mi doctorado. En especial a *Nora Artagaveytia*, por interesarse siempre en mis avances y sacarme todas las dudas de los aspectos clínicos del trabajo.

A la *Cátedra de Dermatología*, por la recolección de las muestras de melanoma y todos los datos clínicos y epidemiológicos.

A *Silvina Bartesaghi*, del CEINBIO, por su gran ayuda con los ensayos en el HPLC. Gracias por tu dedicación y apoyo en la puesta a punto de la técnica y procesamiento de las muestras.

Al *Dr. Manel Esteller*, del Instituto de Investigaciones Biomédicas del Bellvitge (Idibell), por abrirme las puertas de su laboratorio. Y a todos los integrantes del laboratorio que hicieron mis instancias en Barcelona como si estuviera en casa.

A *Javier Carmona* y *Olafur Stefansson*, del Idibell, por cederme gentilmente los datos de metilación obtenidos en tejidos.

A todos los pacientes y personas que participaron de los proyectos de nuestro laboratorio cediendo de manera desinteresada muestras de sangre.

A *mis amigas*, las de toda la vida y las que fui conociendo a lo largo de mi carrera. Son un apoyo constante y siempre están cuando las necesité. En especial a *Claudia*, *Cecilia, Victoria y Alejandra*, quienes comparten conmigo la vocación por la biología y entienden todo lo que ello significa.

A *mis padres y hermanos*, por todo el apoyo que me han dado siempre. Sin ustedes me hubiera sido imposible llegar a donde estoy. No existen las familias perfectas, pero sin duda es la mejor que me pudo tocar.

A *Walter, Mica y Guille* simplemente por quererme y dejar quererlos. Gracias por bancar mi nerviosismo y cansancio, horas de estudio y trabajo en casa, y mis viajes. Prometo reintegrarles cada minuto que no pudimos estar juntos. Walter, gracias por estar siempre a mi lado apoyándome y alentándome a que mejore. Sin dudad nada de esto lo hubiera podido hacer sola.

También quiero agradecer a las siguientes instituciones por el apoyo financiero que me permitieron concluir esta tesis: CHLCC, PEDECIBA, CSIC y ANII.

## INDICE

ABREVIATURAS1
LISTA DE TABLAS Y FIGURAS
RESUMEN
1. INTRODUCCIÓN GENERAL 10
1.1 Epigenética: conceptos generales12
1.1.1 Marcas epigenéticas
Metilación del ADN13
Modificaciones post-traduccionales de histonas
ARN no codificantes
1.1.2 Regulación de la expresión génica de manera cooperativa
1.1.3 Reprogramación epigenética durante el desarrollo embrionario y
mantenimiento de las marcas epigenéticas
Reprogramación epigenética en el embrión temprano
Desregulación epigenética y enfermedades humanas
1.2 PAPEL DE LA METILACIÓN DEL ADN EN EL DESARROLLO TUMORAL
1.2.1 Hipermetilación regional
Inactivación de genes codificantes de proteínas
Hipermetilación de miARNs
"Drivers" y "passengers" epigenéticos
Mecanismos de hipermetilación anormal en cáncer
1.2.2 Hipometilación global
Activación de retrotransposones, oncogenes y miARN 29
Mecanismos involucrados en la hipometilación global del ADN 31
1.2.3 Inactivación mutacional de genes
1.3 EFECTO DE LOS FACTORES AMBIENTALES SOBRE LAS MARCAS EPIGENÉTICAS 33
1.4 Epidemiología epigenética
1.4.1 Variación en la metilación del ADN y enfermedades complejas
1.4.2 Herencia epigenética transgeneracional
1.4.3 Variación epigenética durante el desarrollo 40
1.4.4 Variación epigenética durante el envejecimiento
Deriva epigenética

Pérdida de la plasticidad fenotípica	44
.5 BIOMARCADORES EPIGENÉTICOS EN CÁNCER DE MAMA Y MELANOMA CUTÁNEO	. 45
1.5.1 Cáncer de mama	. 45
Epidemiología y panorama actual en Uruguay	45
Alteración de la metilación del ADN en cáncer de mama	. 46
1.5.2 Melanoma cutáneo maligno	. 47
Epidemiología y panorama actual en Uruguay	. 47
Avances en estudios epigenéticos en melanoma	. 49
1.5.3 Aportes del trabajo de tesis en la búsqueda de biomarcadores de riesgo	en
cáncer de mama y melanoma cutáneo	. 50

2. HIPÓTESIS, OBJETIVOS Y ESTRATEGIA GENERAL DE TRABAJO	51
2.1 Hipótesis de trabajo	51
2.2 Objetivos	51
2.2.1 Objetivo general	51
2.2.2 Objetivos específicos	52
2.3 Estrategia a seguir	52
2.3.1 Abordajes de análisis	52
2.3.2 Población de estudio	53
Estudio en cáncer de mama esporádico	53
Estudio en melanoma cutáneo maligno esporádico	54
Aspectos bioéticos	54
Datos epidemiológicos	55
Muestras biológicas utilizadas	55
Datos de metilación de tejidos y líneas celulares	56

CÁNCER DE MAMA Y MELANOMA ESPORÁDICO	5′
3.1 ANTECEDENTES	5′
3.1.1 Metilación global del ADNgenómico como marcador de riesgo	5′
3.1.2 Variabilidad entre las poblaciones	
3.2 Metodología	6
3.2.1 Cuantificación de metilación global del ADN genómico	6
3.2.2 Genotipado de SNPs y determinación de ancestralidad individual	62
3.2.3 Análisis estadísticos	6

5.5 <b>R</b> ESOLIADOS
3.3.1 Hipometilación del ADN genómico de leucocitos en pacientes con cáncer de
mama y melanoma65
3.3.2 Ancestría genética en la metilación global del ADN genómico de leucocitos
de pacientes con cáncer 69
3.3.3 Asociación de genotipos en genes candidatos y metilación global del ADN
genómico en pacientes con cáncer de mama72
3.4 Discusión
3.4.1 Hipometilación global del ADN de leucocitos como biomarcador de
3.4.1 Hipometilación global del ADN de leucocitos como biomarcador de melanoma y cáncer de mama esporádico
<ul> <li>3.4.1 Hipometilación global del ADN de leucocitos como biomarcador de melanoma y cáncer de mama esporádico</li></ul>
<ul> <li>3.4.1 Hipometilación global del ADN de leucocitos como biomarcador de melanoma y cáncer de mama esporádico</li></ul>
<ul> <li>3.4.1 Hipometilación global del ADN de leucocitos como biomarcador de melanoma y cáncer de mama esporádico</li></ul>

# 4. PATRONES DE METILACIÓN DEL ADN EN MELANOMA CUTÁNEO

ESPORÁDICO
4.1 Antecedentes
4.1.1 Marcadores de riesgo en melanoma cutáneo
4.1.2 Alteraciones de la metilación del ADN asociadas a pronóstico y progresión
en melanoma 88
4.2 Metodología
4.2.1 Muestras de ADN utilizadas en el estudio
4.2.2 Amplificación entre sitios metilados (AIMS) 89
4.2.3 Visualización de los resultados de AIMS
Electroforesis en gel de secuencia desnaturalizante
Tinción con Nitrato de Plata92
Interpretación de los resultados de AIMS
4.2.4 Caracterización de fragmentos de ADN con metilación diferencial
4.2.5 Validación de las regiones diferencialmente metiladas identificadas por
AIMS en una cohorte muestral independiente
4.2.6 Análisis estadísticos
4.3 Resultados
4.3.1 Perfiles de metilación del ADN de pacientes con melanoma cutáneo e
individuos no afectados97

4.3.2 Caracterización de los fragmentos de ADN con metilación diferencial entre
pacientes y controles101
4.3.3 Análisis en tejidos y líneas celulares de las MVPs detectada en los genes
NTM y TMEM184A 103
Región de banda A1 (gen NTM)103
Región de banda B2 (gen TMEM184A) 105
4.3.4 Selección de sitios CpG candidatos para su futura validación en muestreo
independiente correspondiente a nuestra población
4.4 DISCUSIÓN
4.4.1 Aplicación de la técnica de AIMS en estudios poblacionales108
4.4.2 Regiones con metilación diferencial del ADN de leucocitos en pacientes con
melanoma cutáneo111
Neurotrimina112
TMEM184A 115
Siguiente paso en el estudio de NTM y TMEM184A como
biomarcadores116
4.4.3 Conclusiones

# 

PORADICO	
5.1 ANTECE	EDENTES
5.2 Metodo	0LOGÍA
5.2.1	Descripción de la población de estudio120
5.2.2	Infinium HumanMethylation450 BeadChip120
5.2.3	Identificación de CpG diferencialmente metilados entre pacientes y controles
5.2.4	Anotación génica, análisis de ontología génica y enriquecimiento de vías
	moleculares
5.2.5	Análisis de bases de datos de expresión génica
5.2.6	Validación en muestras de tejidos y líneas celulares
5.2.7	Secuenciación de ADN tratado con bisulfito124
5.3 Result	ADOS
5.3.1	Perfiles de metilación del ADN de leucocitos126
5.3.2	Los análisis de agrupamientos no supervisados revelan un perfil de
	metilación del ADN específico de pacientes con cáncer de mama 129
5.3.3	Descripción de los CpGDM detectados

Análisis in sílico de datos de expresión génica disponibles públicamente
5.3.4 Validación del panel de CpGDMs en tejidos de muestreo independiente de
origen europeo139
5.3.5 Selección de CpGDM candidatos para su validación clínica 142
5.4 DISCUSIÓN
5.4.1 Perfil de metilación diferencial en pacientes con cáncer de mama 145
5.4.2 Sitios CpGDM seleccionados para su validación técnica y clínica 148
5.4.3 Conclusiones

## 

# 

8. ANI	EXOS	176
	8.1 TABLAS SUPLEMENTARIAS	176
	8.2 FIGURAS SUPLEMENTARIAS	183

#### ABREVIATURAS

0 <sup>6</sup> -MGMT	O <sup>6</sup> -metilguanina DNA metiltransferasa
5hmC	5-hidroximetilcitosina
5mC	5-metilcitosina
5mdC	5-metil-2'-deoxicitidina
ADN	Ácido desoxirribonucleico
AIM	Marcador informativo de ancestralidad
AIMS	Amplificación entre sitios metilados
ARN	Ácido ribonucleico
ARNm	ARN mensajero
ARNnc	ARN no codificante
ARNpol	ARN polimerasa
ASM	Metilación alelo específica
CART	Árbol de clasificación y regresión
CHR	Cromosoma
CpG	Dinucleótido citosina-guanina
CpGDM	CpG diferencialmente metilado
CpNpG	Trinucleótido citosina-cualquier base-guanina
dA	Deoxiadenina
dC	Deoxicitidina
dG	Deoxiguanosina
dNTPs	Desoxirribonucleótidos trifostatos
dT	Deoxitimidina
dU	Deoxiuridina
DMR	Región diferencialmente metilada
DNMT	Metiltransferasa de ADN
EWAS	Estudio de asociación epigenómico
FDR	Tasa de falso descubrimiento

GWAS	Estudio de asociación genómico
H1	Histona tipo 1
H2A	Histona tipo 2A
H2B	Histona tipo 2B
Н3	Histona tipo 3
H4	Histona tipo 4
HAT	Acetiltransferasa de histonas
HDAC	Desacetilasa de histonas
HDMT	Desmetilasa de histonas
HERV-K	Retrovirus endógeno humano K
НМТ	Metiltransferasa de histonas
HNPCC	Cáncer colorrectal no polipósico hereditario
HPLC	Cromatografía líquida de alta resolución
HSM	Metilación haplotipo específica
ICM	Masa celular interna
IMC	Índice de masa corporal
КС	Control no afectado con cáncer
KP	Paciente con cáncer de mama
LC	Línea celular
LINE	long interspersed nucleotide element-1
lncARN	ARN no codificante largo
MBD	Proteína de unión a CpG metilado
MC	Control no afectado con melanoma
MDS	Análisis de escalamiento multidimensional
metQTLs	methylation quentitative trait loci
mH3K27	Histona H3 trimetilada en la lisina 27 del extremo N-terminal
mH3K36	Histona H3 trimetilada en la lisina 36 del extremo N-terminal
mH3K4	Histona H3 trimetilada en la lisina 4 del extremo N-terminal
mH3K79	Histona H3 trimetilada en la lisina 79 del extremo N-terminal

mH3K9	Histona H3 trimetilada en la lisina 9 del extremo N-terminal
mH4K20	Histona H3 trimetilada en la lisina 20 del extremo N-terminal
miARN	microARN
MM	Paciente con melanoma maligno cutáneo
MS-HRM	Análisis de curvas de fusión de alta resolución sensible a la metilación
MVP	Posición de metilación variable
MZ	Gemelos monocigóticos
NGS	Secuenciación masiva de nueva generación
PBL	Leucocitos de sangre periférica
PBMC	Células mononucleares de sangre periférica
PCR	Reacción en cadena de la polimerasa
PGC	Células germinales primordiales
RE	Receptor de estrógeno
RP	Receptor de progesterona
SAM	S-adenosil-L-metionina
siARN	ARN pequeño de interferencia
SINE	Short interspersed nuclear element
SNP	Polimorfismo de nucleótido simple
TF	Factor de transcripción
TMN	Tejido de mama normal
TP	Tumor primario de mama
TSS	Sitio de inicio de la transcripción
UTR	Región no traducida

#### LISTA DE TABLAS Y FIGURAS

**Figura 1.** Ilustración esquemática de la metilación del ADN catalizada por las ADN metiltransferasas (DNMTs) utilizando S-adenosil-L-metionina como donador del grupo metilo.

**Figura 2.** Patrones de metilación del ADN en diferentes regiones del genoma y consecuencias en la regulación. a) Islas CpG en promotores génicos; b) CpG shores, localizado a 2 kb de las islas CpG; c) Secuencias repetidas.

**Figura 3.** Metilación del ADN en regiones intragénicas. A) Regulación de promotores intragénicos alternativos. B) Regulación de la actividad en potenciadores. C) Regulación génica via splicing alternativo. D) Regulación de ARN no codificantes (ARNnc) intragénicos.

**Figura 4.** Modificaciones de histonas. Las colas de todas las histonas son susceptibles de modificaciones post-traduccionales en sus colas.

**Figura 5.** Mecanismo mediante el cual la metilación del ADN y la desacetilacón de histonas cooperan para reprimir la transcripción.

**Figura 6.** El cáncer es el resultado de la combinación de alteraciones genéticas y epigenéticas.

Figura 7. Interacción dinámica entre genotipo, epigenotipo y fenotipo.

**Figura 8.** Diferentes tipos de variación en la metilación del ADN que pueden ser identificados con EWAS (estudios de asociación epigenéticos).

**Figura 9.** Gráfico de Forest de asociación entre riesgos a cáncer e hipometilación global del ADN en leucocitos de sangre periférica.

**Figura 10.** Separación de los deoxinucleósidos por HPLC. Cromatogramas obtenidos por detección UV a una absorbancia a 254 nm y 280 nm.

**Figura 11.** Comparación de los niveles de metilación global del ADN genómico en leucocitos de pacientes con cáncer e individuos controles.

**Figura 12.** Distribución de los niveles de metilación global del ADN en cada individuo del estudio. Frecuencia de los niveles de metilación a lo largo de (A) 86 pacientes con cáncer de mama y 92 controles, y (B) 42 pacientes con melanoma cutáneo y 46 controles.

**Figura 13.** Correlación entre la ancestría genética individual y la metilación global del ADN en leucocitos de pacientes con cáncer. (a) El componente ancestral africano se correlaciona negativamente con la metilación del ADN (r = -0.187, p<0.005). (b) El

componente ancestral europeo se correlaciona positivamente con la metilación del ADN (r = 0.169, p<0.01).

**Figura 14.** Análisis de árboles de clasificación y regresión (CART) de metilación global del ADN de leucocitos y componente genético ancestral en pacientes con cáncer.

Figura 15. Diagrama de la técnica de AIMS.

**Figura 16.** Purificación de banda de interés mediante varias rondas de extracción de la banda del gel y reamplificación. A: Gel acrilamida 6%/urea 8M luego de la primera purificación y reamplificación. B: Gel acrilamida 6%/urea 8M luego de la segunda ronda de purificación y reamplificación. C: Gel agarosa 2% luego de la última purificación.

**Figura 17.** *Fingerprinting* de metilación obtenido mediante AIMS y visualizado en gel de poliacrilamida desnaturalizante.

**Figura 18.** Ejemplo de perfiles de metilación obtenidos por AIMS con la reacción PCR-A en muestras de ADN de leucocitos de pacientes con melanoma y grupo control.

**Figura 19.** Dendograma basado en agrupamiento jerárquico de los pacientes con melanoma y controles no afectados, considerando los perfiles de metilación de las 9 bandas variables obtenidas por la reacción PCR-A de AIMS.

**Figura 20.** Localización cromosómica del fragmento correspondiente a la banda A1 diferencialmente metilada, aislada de los perfiles de AIMS. A) Localización de la secuencia A1 en el contexto de las regiones génicas codificantes cercanas e islas CpG identificadas en Genome Browser. B) Ampliación del fragmento de 790 pb correspondiente a la banda A1 aislada.

**Figura 21.** Localización de la secuencia B2 en el contexto de las regiones génicas codificantes cercanas e islas CpG identificadas en Genome Browser.

**Figura 22.** Sondas evaluadas por la plataforma 450K HumanMethylation BeadChip (Illumina) dentro y en la cercanía de la secuencia de la banda A1 aislada por AIMS.

**Figura 23.** Comparación de los niveles de metilación de sitios CpG en la región de la secuencia A1 en tejidos de nevos normales, tumores primarios de melanoma y tumores metastásicos, analizados en la plataforma Human Methylation450 BeadChip. A) Significancia estadística (p valor) de las diferencias observadas entre los distintos tejidos por prueba Wilcoxon. B) Gráfico del promedio de nivel de metilación (valor  $\beta$ ) de cada grupo de tejidos para cada sitio CpG.

**Figura 24.** Sondas evaluadas por la plataforma 450K HumanMethylation BeadChip (Illumina) dentro y en la cercanía de la secuencia de la banda B2 aislada por AIMS.

**Figura 25.** Comparación de los niveles de metilación de sitios CpG en la región de la secuencia B2 en tejidos de nevos normales, tumores primarios de melanoma y tumores metastásicos y líneas celulares, analizados en la plataforma HumanMethylation450 BeadChip. A) Significancia estadística (p valor) de las diferencias observadas entre los distintos tejidos por prueba Wilcoxon. B) Gráfico del promedio de nivel de metilación (valor  $\beta$ ) de cada grupo de tejidos para cada sitio CpG.

**Figura 26.** Distribución de frecuencia de los niveles de metilación de los sitios CpG representados en el array 450K HumanMethylation BeadChip en el ADN de leucocitos de los 32 individuos analizados.

**Figura 27.** Distribución de frecuencia de los niveles de metilación de los sitios CpG representados en el array 450K HumanMethylation BeadChip en el ADN de leucocitos de: (a) controles no afectados y (b) pacientes con cáncer de mama.

**Figura 28.** Nivel de metilación de los sitios CpGs identificados por el ensayo Infinium 450K HumanMethylation.

**Figura 29.** Agrupamiento jerárquico de 45.000 CpGs elegidos al azar en 22 pacientes con cáncer de mama (KP) y 10 controles sanos (KC) analizados en la plataforma Infinium HumanMethylation450 BeadChip.

**Figura 30.** Análisis de escalamiento multidimensional (MDS) de los valores de metilación de los CpGDMs en leucocitos de 22 pacientes con cáncer de mama y 10 controles no afectadas.

**Figura 31.** Agrupamiento jerárquico de los CpGDM en 22 pacientes con cáncer de mama (KP) y 10 controles sanos (KC) analizados en la plataforma Infinium HumanMethylation 450K BeadChip.

**Figura 32.** Dendograma basado en agrupamiento jerárquico mediante boostrap de 10.000 remuestreos, para los valores de metilación de los 77 CpGDM en pacientes con cáncer de mama (KP) y controles (KC).

**Figura 33.** Diferencias en la media de los valores betas entre pacientes con cáncer de mama y controles para los 77 sitios CpG diferencialmente metilados identificados por prueba de Wilcoxon (p<0.05).

**Figura 34.** Distribución genómica de los 77 CpGDMs de acuerdo a su respectiva localización génica (a) y contexto CpG (b).

**Figura 35.** Comparación de los CpGDMs con la distribución genómica y contexto CpG esperado de acuerdo a los CpG impresos en el 450K HumanMethylation BeadChip.

**Figura 36.** Agrupamiento en dendograma con remuestreo boostrap (n = 10.000) de los 77 CpGDMs entre muestras de tejido mamario sano (TMN), tumores primarios de mama (TP) y líneas celulares de mama (LC).

**Figura 37.** Agrupamiento jerárquico de los 77 CpGDM en 14 muestras pareadas de tejido mamario normal (N) y tumor primario (T) analizados en la plataforma Infinium HumanMethylation450 BeadChip.

**Figura 38.** Secuenciación con bisulfito de los clones individuales a las regiones flanqueantes al sitio CpG evaluado por la sonda cg2658226 en el gen *CYFIP1*.

**Tabla 1.** Características epidemiológicas evaluadas en los pacientes con cáncer y los individuos controles con datos de metilación global del ADN.

**Tabla 2.** Media de la metilación global del ADN genómico de leucocitos en pacientes con cáncer y controles.

**Tabla 3.** Comparación del componente genético ancestral individual entre pacientes con cáncer y controles no afectados.

**Tabla 4.** Coeficientes de correlación de Kendall y sus significancias estadísticas entre metilación global del ADN de leucocitos, edad y componentes individuales ancestrales.

**Tabla 5.** Asociación entre SNPs en genes candidatos de cáncer y niveles de metilación genómica global en leucocitos de individuos del estudio caso-control de cáncer de mama.

**Tabla 6.** Comparación de la presencia de las bandas variables entre pacientes con melanoma y controles no afectados.

Tabla 7. Cebadores utilizados en la secuenciación del ADN tratado con bisulfito.

**Tabla 8.** Características de los 77 CpG diferencialmente metiladas entre pacientes con cáncer de mama y grupo control no afectado.

**Tabla 9.** CpGDMs en genes previamente asociados a cáncer en Cancer Gene Census y

 G2SBC Database.

**Tabla 10.** Genes con CpGDMs que presentan diferencias significativas de expresión en PBMC de pacientes con cáncer de mama e individuos control con mamografías normales (Estudio GSE27562).

**Tabla 11.** Meta-análisis de expresión génica en muestras de tumores de mama vs tejido mamario sano.

**Figura suplementaria 1.** Agrupamiento jerárquico basado en los valores de metilación de los 77 CpGDM de muestras de tumores primarios mamarios (TP) y tejidos mamarios sanos (TMN).

**Figura suplementaria 2.** Secuenciación con bisulfito de los clones individuales a las regiones flanqueantes al sitio CpG evaluado por la sonda cg14024502 en el gen *MAP3K6*.

**Figura suplementaria 3.** Secuenciación con bisulfito de los clones individuales a las regiones flanqueantes a los sitios CpGs evaluados por las sondas cg19246761 y cg01229567 en el gen *MIB2*.

**Tabla Suplementaria 1.** Líneas celulares comerciales utilizadas del banco de tejidosdel Programa de epigenética y biología del cáncer (PEBC) del IDIBELL.

**Tabla Suplementaria 2**. Lista de marcadores informativos de ancestralidad (MIAs) utilizados para determinar la ancestría genética individual, con la posición cromosómica y los alelos.

**Tabla Suplementaria 3.** Modelo predictivo de regresión de valores promedio de metilación global del ADN de acuerdo a los genotipos del rs545500 en gen *MUS81*.

**Tabla Suplementaria 4.** Características epidemiológicas y tumorales de las pacientes con cáncer de mama analizadas en la plataforma 450K HumanMethylation BeadChip.

**Tabla Suplementaria 5.** Características de los tumores correspondientes a los pares de tejidos mamarios (normal/tumoral) de pacientes con cáncer de mama.

**Tabla Suplementaria 6.** Comparación de la distribución genómica de los CpGDM con respecto a la representación de las CpG en la plataforma 450K HumanMethylation BeadChip.

**Tabla Suplementaria 7.** Comparación de los perfiles de metilación de los 77 CpGDMs en los distintos tipos de muestras analizadas.

**Tabla Suplementaria 8.** Comparación de los valores de metilación de los 77 CpGDMentre muestras de ADN de tumores primarios de mama y tejidos mamarios sanos.

**Tabla Suplementaria 9.** Características asociadas a los criterios de selecciónestablecidos de los CpGDM candidatos.

#### RESUMEN

El estudio únicamente de variantes genéticas no es suficiente para explicar una enfermedad compleja como el cáncer. Alteraciones epigenéticas, entre ellas la metilación del ADN, han sido asociadas con diferentes tipos de tumores.

Con el fin de detectar marcadores epigenéticos de riesgo en sangre para el desarrollo de cáncer, buscamos detectar metilación diferencial en el ADN de leucocitos de una muestra de pacientes con melanoma cutáneo y cáncer de mama esporádico mediante dos acercamientos metodológicos complementarios. En primer lugar, cuantificamos el nivel global de citosinas metiladas por cromatografía líquida de alta resolución (HPLC-ESI/MS). Confirmada la existencia de diferencias en los niveles genómicos de metilación entre pacientes y controles, realizamos estudios para localizar las regiones genómicas con metilación diferencial sitio-específica empleando un análisis sesgado de una porción del genoma (Amplificación entre sitios metilados), y un análisis de alta resolución a nivel de sitios CpG basado en tecnología de microarreglos.

Como resultado de estos estudios, demostramos la ocurrencia de hipometilación global del ADN genómico de leucocitos en pacientes con cáncer de mama y melanoma esporádico, y el efecto del contexto genético sobre la variabilidad epigenética en nuestra población, dado por la ancestría genética de los individuos y/o variantes genéticas en genes asociados directa o indirectamente con los mecanismos epigenéticos.

A nivel de metilación específica, identificamos dos regiones diferencialmente metiladas en leucocitos de pacientes con melanoma cutáneo en comparación a individuos sanos, ubicadas en los genes Neurotrimina y TMEM184A. Dichas regiones fueron caracterizadas y evaluadas en muestras de tejidos de individuos con melanoma cutáneo de origen europeo.

Por último, mediante el análisis de más de 485.000 sitios CpG detectamos un panel de 77 sitios CpGs diferencialmente metilados en sangre de pacientes con cáncer de mama, que incluye los genes CYFIP1, MAP3K6 y MIB2, el cual fue validado en un muestreo independiente de tejidos de mama pertenecientes a pacientes con cáncer de mama de origen europeo. Este panel no solo diferencia leucocitos de pacientes con cáncer de mama esporádico de controles, sino que también tiene la potencialidad de diferenciar tejido mamario sano del tumoral incluso en casos de cáncer de mama hereditario.

En este trabajo se describe por primera vez un grupo de marcadores epigenéticos candidatos en leucocitos de sangre periférica asociados a tumores sólidos en la población uruguaya. La validación de estos resultados en un muestreo mayor e independiente en un futuro próximo, abrirá las puertas a la evaluación de paneles de biomarcadores propios de nuestra población.

# **Introducción general**

Las alteraciones genéticas pueden explicar las causas de varias enfermedades monogénicas. Sin embargo, la base genética subyacente en el origen de las enfermedades complejas y multifactoriales sigue siendo en gran medida desconocida, y el papel de los mecanismos no genéticos, incluidos los mecanismos epigenéticos o modificaciones postraduccionales de proteínas, ha tomado gran importancia.

En la etiología de las enfermedades complejas, la epigenética puede ayudar a proveer una explicación de tres de sus características: i) la dependencia de la mayoría de las enfermedades complejas con la edad, la cual no se puede explicar muy bien solo por la acumulación de mutaciones; ii) su naturaleza cuantitativa; y iii) el mecanismo por el cual el ambiente puede modular la predisposición genética a la enfermedad (Bjornsson et al., 2004).

El cáncer es la enfermedad humana compleja asociada con defectos epigenéticos mejor caracterizada (Berdasco and Esteller, 2010). Los cambios epigenéticos ocurren en todos los tipos de cáncer e incluyen alteraciones en la metilación del ADN, modificaciones de histonas y cambios de expresión de ARNs no codificantes (ARNnc). Hoy en día, existe suficiente evidencia que las alteraciones epigenéticas y sus variaciones interindividuales influyen sobre el riesgo a cáncer y pueden emplearse como biomarcadores potenciales.

Como describiremos en detalle a continuación, es claro que los mecanismos epigenéticos puede jugar un papel fundamental en mediar los efectos de medioambiente, y existen varios mecanismos por los cuales factores genéticos y epigenéticos interactúan. Así, el campo de la epidemiología epigenética del cáncer tiene un gran reto que es elucidar las interacciones genética-epigenética-medioambiente, y determinar la forma en que modulan el riesgo de cáncer individual (Risch et al., 2012).

En el pasado, los estudios epigenéticos se focalizaban fundamentalmente en analizar pares de tejido sano/tumoral. Sin embargo, una pregunta importante para la epidemiología del cáncer es si los cambios epigenéticos pueden ser identificados en tejidos no afectados directamente por la patología, como ADN aislado de linfocitos periféricos, muestras de plasma, orina, o hisopado bucal. La detección de cambios epigenéticos en tejidos no tumoral y de fácil acceso mediante técnicas mínimamente invasivas aumentaría en gran medida el potencial marcador de dichas alteraciones.

En los futuros estudios será importante caracterizar la interacción entre factores ambientales, estilos de vida, factores genéticos y epigenéticos. Así, los abordajes integrando la información genética, epigenética, epidemiológica y clínica de los individuos son los más prometedores y podrían mejorar los modelos de riesgo para el cáncer actualmente disponibles.

## 1.1 Epigenética: conceptos generales

Más de una década ha pasado desde que el genoma humano fuera completamente secuenciado, pero aún queda por ser elucidado cómo la información genómica dirige los programas de expresión génica espacial y temporal-específica (Lander, 2011). La respuesta a esta pregunta no solo es esencial para comprender los mecanismos de desarrollo embrionario humano sino también para estudiar las variaciones fenotípicas entre las poblaciones humanas y la etiología de muchas enfermedades. Sin embargo, el mayor desafío sigue siendo el siguiente: cada uno de las más de 200 tipos celulares diferentes en el cuerpo humano contiene una copia idéntica del genoma pero expresa un grupo de genes particular. ¿Cómo el genoma guía a un grupo limitado de genes a ser expresados a diferentes niveles en los distintos tipos celulares? Hoy en día, hay grandes evidencias que indican que el epigenoma determina en gran medida el programa de expresión génica en cada tipo celular, junto con su genoma.

La palabra "epigenética", surgida hace más de medio siglo, combinando "epigénesis" y "genética", describe los mecanismos de compromiso de destino celular y especificación de linaje durante el desarrollo animal (Holliday, 1990; Waddington, 1959). Hoy, "epigenoma" es generalmente utilizado para describir el conjunto de procesos que modulan en una célula la estructura de la cromatina, y con ello sus funciones (expresión génica, reparación del ADN, recombinación), independiente de la secuencia de ADN (Bird, 2007). Estos mecanismos epigenéticos involucran la metilación del ADN y las modificaciones covalentes de las historias a lo largo del genoma, así como la acción de ARN no codificantes (Bonasio et al., 2010). El epigenoma puede diferir de un tipo celular a otro, y en cada célula regula la expresión génica de varias maneras: organizando la estructura nuclear de los cromosomas, restringiendo o facilitando el acceso de factores de transcripción al ADN, y preservando la memoria de actividades transcripcionales pasadas. De esta manera, el epigenoma puede ser considerado una segunda dimensión de la secuencia de ADN y es crucial para el mantenimiento de los patrones de expresión génica específico del tipo celular (Rivera and Ren, 2013).

A diferencia de los alelos genéticos, los epi-alelos no difieren en su secuencia de ADN; la información epigenética reside en firmas moleculares, que pueden surgir *de* 

*novo*, y que se transmiten durante la división celular y proporcionan un recuerdo de los estímulos previamente experimentados, sin cambios irreversibles en la información genética (Bonasio et al., 2010).

Es habitual que la metilación del ADN, las modificaciones de histonas y los ARNnc sean descritos como mecanismos independientes, pero es importante notar que existe una interacción muy estrecha entre las diferentes marcas epigenéticas para regular el estado de la cromatina y de la expresión génica. El proyecto ENCODE (ENCODE Project Consortium), un gran esfuerzo colaborativo desarrollado para definir todos los elementos funcionales en el genoma humano, recientemente ha publicado grandes sets de datos de transcripción, modificaciones de histonas y proteínas de unión. Estos datos describen perfiles epigenéticos superpuestos tanto regional como globalmente, los cuales se combinan para regular la expresión génica (Bernstein et al., 2012; ENCODE Project Consortium., 2004).

#### 1.1.1 Marcas epigenéticas

#### Metilación del ADN

La metilación del ADN es un proceso epigenético involucrado en muchos procesos celulares incluyendo la expresión génica, impronta genómica, inactivación del cromosoma X, silenciamiento de retrovirus y elementos transponibles, y organización de la cromatina.

La metilación del ADN consiste en la adición covalente post-replicación de un grupo metilo (-CH<sub>3</sub>) en el carbono 5 del anillo de la citosina para generar 5metilcitosina (5mC). Esta reacción enzimática es catalizada por una familia de enzimas ADN metiltransferasas (DNMTs) que transfieren un grupo metilo desde la S-adenosil-L-metionina (SAM) (Figura 1). La adición del grupo metilo en la citosina no interfiere estéricamente con el apareamiento de bases, pero si interfiere con la interacción de algunas proteínas de unión al ADN debido a que se altera la estructura de la cromatina. Citosinas metiladas están presentes en el ADN de todos los vertebrados y plantas, algunos hongos e invertebrados, y muchas especies de bacterias (Bird, 2002).



**Figura 1.** Ilustración esquemática de la metilación del ADN catalizada por las ADN metiltransferasas (DNMTs) utilizando S-adenosil-L-metionina como donador del grupo metilo (Modificado de Khan et al. 2012).

En mamíferos se han detectado 5 miembros de la familia DNMTs: DNMT1, DNMT2, DNMT3A, DNMT3B y DNMT3L, pero solo DNMT1, DNMT3A y DNMT3B poseen actividad metiltransferasa. Estas enzimas se clasifican en: *DNTMs de mantenimiento* (DNTM1) que preservan los patrones de metilación del ADN durante la división celular, actuando específicamente sobre el ADN hemimetilado; y DNMTs *de novo* que establecen los patrones de metilación en secuencias específicas durante el desarrollo embrionario (DNMT3A y DNMT3B) (Jin and Robertson, 2013). Las DNMTs *de novo* tienen una alta expresión en células madres embrionarias y luego disminuye su expresión en las células diferenciadas. DNMT3L, a pesar de ser catalíticamente inactiva, es necesaria para el establecimiento del *imprinting* genómico durante la gametogénesis (Bourc'his et al., 2001) y actúa como factor estimulador de DNMT3A y DNMT3B. La DNMT1 es la DNMT más abundante en la célula y se expresa mayormente durante la fase S del ciclo celular. Tiene alta preferencia por el ADN hemimetilado que se generan durante la replicación del ADN semiconservativa (Denis et al., 2011).

La metilación podría tener lugar en cualquier residuo de citosina del genoma. Sin embargo, la transmisión clonal de los patrones de metilación está restringida a citosinas de secuencias simétricas CpG y CpNpG (N hace referencia a cualquier nucleótido), es decir, en secuencias cuya lectura es palindrómica. La metilación que tiene lugar en trinucleótidos CpNpG se observa fundamentalmente en plantas. Los dinucleótidos CpG tienden a agruparse en regiones denominadas "islas CpG", definidas como regiones de más de 200 pb con alto contenido en secuencias CG y con relación frecuencia observada/esperada de CpG > 0.6 (Portela and Esteller, 2010).

Hasta recientemente, la mayoría de los estudios de metilación del ADN se focalizaron en el análisis de islas CpG asociadas a secuencias reguladoras. Sin embargo, con el avance del estudio de la metilación del ADN a nivel genómico, como veremos a continuación, se ha demostrado que no sólo la metilación en promotores está modulada durante los procesos fisiológicos y las enfermedades, sino también las regiones intragénicas e intergénicas (Irizarry et al., 2009).

Los dinucleótidos CpG son usualmente raros en los genomas de mamíferos (~1%). Sin embargo, el ADN de los tejidos somáticos de mamíferos está metilado en el 70% de los dinucleótidos CpGs. Las secuencias altamente metiladas incluyen ADN satélite, elementos repetidos como transposones, ADN intergénico no repetido y exones de genes. Evidencias experimentales indican que la metilación del ADN sirve como un mecanismo de defensa para silenciar elementos móviles o invasivos del genoma, virales. impidiendo la inestabilidad transposones y secuencias genómica, translocaciones e inactivación génica por reactivación de secuencias endoparásitas (Figura 2) (Eden et al., 2003). La excepción a esta metilación global del genoma de mamíferos son las islas CpGs, regiones que habitualmente están sin metilar en las células germinales, en el embrión temprano y generalmente también en todos los tejidos somáticos (Bird, 2002). La mayoría de estas islas se solapan con los promotores génicos, los potenciadores y elementos silenciadores, dando lugar a una estructura abierta de la cromatina permitiendo la expresión génica por encontrarse desmetiladas. Alrededor del 60% de los genes codificadores de proteínas en mamíferos presentan islas CpG en sus regiones promotoras y usualmente están desmetilados en células normales, aunque algunos promotores (~6%) se metilan de manera tejido-específico durante el desarrollo temprano o en tejidos diferenciados (Straussman et al., 2009) (Figura 2).



**Figura 2.** Patrones de metilación del ADN en diferentes regiones del genoma y consecuencias en la regulación. a) Islas CpG en promotores génicos; b) CpG shores, localizado a 2 kb de las islas CpG; c) Secuencias repetidas. Modificado de Portela & Esteller 2010.

Recientemente ha surgido el término CpG shores, refiriéndose a regiones de menor densidad CpG ubicadas a gran proximidad (~2 kb) de las islas CpG. La metilación en estas regiones está asociada con inactivación transcripcional (Figura 2). La mayoría de la metilación del ADN tejido-específica parece que ocurre no en islas CpG sino en CpGshores (Doi et al., 2009; Irizarry et al., 2009).

En los últimos años se ha demostrado la importancia de la metilación de CpG en el cuerpo génico o "*gene body*" sobre la expresión génica, entendiendo como cuerpo génico los exones (excluyendo el exón 1), intrones y región 3′ no traducida (Lister et al., 2009). La metilación intragénica en el genoma humano tiene un potencial papel en el uso de promotores alternativos, regulación de ARNnc, procesamiento alternativo del ARN así como en la actividad de potenciadores (revisado en Kulis et al. 2013) (Figura 3).

Aunque la metilación del ADN ocurre principalmente en un contexto de dinucleótidos CpG en mamíferos, se ha descrito recientemente en humanos metilación en sitios no CG, específicamente en sitios CHG y CHH (donde H es A, C o T) (Lee et al., 2010). La metilación en CHG y CHH ha sido encontrada en células madres y parecería estar enriquecida en los cuerpos génicos y ausente en sitios de unión de proteínas y potenciadores (Lister et al., 2009). Los niveles de la metilación no-CpG disminuyen durante la diferenciación y son restaurados en las células madres pluripotentes inducidas, sugiriendo un rol clave en el origen y mantenimiento del estado

pluripotente; aunque los mecanismos de metilación no-CpG aún no son claros (Laurent et al., 2010; Lister et al., 2009, 2011).



Figura 3. Metilación del ADN en regiones intragénicas. A) Regulación de promotores intragénicos alternativos. La metilación de un promotor alternativo (P2) resulta en el silenciamiento de este promotor por unión de proteínas de unión al ADN metilado (MBD) o por bloqueo de factores de transcripción (TF). La desmetilación resulta en la activación del promotor P2, facilitando la unión de TFs o mediante acción de un potenciador distal (indicado en azul). B) Regulación de la actividad en potenciadores. La metilación de un potenciador intragénico (indicado en azul) resulta en inactivación transcripcional de su correspondiente gen. La desmetilación de este potenciador resulta en la unión de TFs, el curvado del ADN y/o expresión de un ARN no codificante regulador (eRNA). C) Regulación génica vía splicing alternativo. La metilación en un exón alternativo (E2) bloquea la unión de la proteína CTCF y la ARNpol II no permite que el spliceosoma se ensamble y el E2 es salteado. La desmetilación del sitio de splicing en E2 resulta de la unión de CTCF, pausando la ARNpol II, el ensamblado del spliceosoma y por lo tanto la inclusión de E2 en el ARNm. D) Regulación de ARN no codificantes (ARNnc) intragénicos. La desmetilación de promotores de ARNnc resulta en su expresión y regulación de los genes blancos. La regulación puede ser en cis, afectando el gen huésped o en trans, regulando genes distantes. Los círculos negros y blancos representan CpG metilados y no metilados, respectivamente. Modificado de Kulis et al. 2013.

A pesar de que la metilación del ADN es el proceso epigenético más estudiado, una de las preguntas más intrigantes aún no respondida es cómo la maquinaria de metilación es dirigida a secuencias específicas en el genoma. Se han propuestos varios mecanismos, principalmente sugiriendo la interacción de DNMTs con otros factores epigenéticos. Más recientemente se ha descrito en plantas la metilación del ADN dirigida por ARN, en la que participarían pequeños ARNs inhibitorios (siRNA) (He et al., 2011). Aunque este proceso está bien estudiado en plantas y algunos de los componentes de la metilación del ADN dirigida por ARN están conservados en

mamíferos, no está claro si en animales están involucrados procesos similares en la regulación de la metilación del ADN (He et al., 2011).

#### Modificaciones post-traduccionales de histonas

El segundo jugador clave en la conformación de la cromatina y la regulación génica son las modificaciones covalentes de las colas de las histonas. Las histonas H2A, H2B, H3 y H4 se agrupan en dos dímeros H2A-H2B y un tetrámero H3-H4 para formar el nucleosoma. Un segmento de 147 pb de ADN se enrolla en 1.65 vueltas alrededor del octámero de histonas y los nucleosomas vecinos están separados por, en promedio, ~50 pb de ADN libre. Las histonas nucleosómicas son predominantemente globulares excepto por sus colas N-terminales, las cuales no tienen estructura secundaria. La histona H1 se denomina histona linker, ya que no forma parte del nucleosoma pero mantiene al octámero en su lugar uniéndose al ADN en el sitio de entrada y salida del nucleosoma. Las colas de todas las histonas son susceptibles de modificaciones posttraduccionales: acetilación, metilación, fosforilación, ubiquitinación, sumolación y ADP ribosilación, entre otras (Figura 4). Las modificaciones de las histonas tienen importantes roles en la regulación de la expresión génica, reparación del ADN, replicación del ADN, splicing alternativo y condensación de los cromosomas (Kouzarides, 2007).



**Figura 4.** Modificaciones de histonas. Las colas de todas las histonas son susceptibles de modificaciones post-traduccionales en sus colas. Se marcan las modificaciones principales: acetilación (azul), metilación (rojo), fosforilación (amarillo) y ubiquitinación (verde). El número en gris bajo cada aminoácido representa la posición en la secuencia (Portela and Esteller, 2010).

En relación a su estado transcripcional, el genoma humano puede ser dividido en eucromatina activa y heterocromatina inactiva. La eucromatina está caracterizada por altos niveles de acetilación en las histonas del nucleosoma, así como por trimetilación en la lisina 4 de la histona H3 (mH3K4), lisia 36 de la histona H3 (mH3K36) y lisina 79 de la histona H3 (mH3K79). Por otra parte, la heterocromatina es caracterizada por bajos niveles de acetilación de histonas y altos niveles de metilación en lisina 9 de la histona H3 (mH3K9), lisina 27 de la histona H3 (mH3K27) y lisina 20 de la histona H4 (mH4K20) (Li et al., 2007). Así, la acetilación de histonas está asociada con la activación de la transcripción debido a que el grupo acetilo neutraliza las cargas positivas de las histonas, reduce la afinidad por el ADN y contribuye a la expansión de la estructura de la cromatina. Sin embargo, la metilación de las histonas tiene como resultado varias consecuencias transcripcionales dependiendo qué histona es afectada, qué residuo es modificado y el número de grupos químicos que se adicionan a cada residuo.

Las histonas pueden ser modificadas en diferentes sitios simultáneamente, y existe una comunicación entre las marcas que ocurren en la misma cola de la histona y en colas de diferentes histonas. Así, una única marca no determina un cambio en la cromatina; pero, la combinación de todas las marcas en un nucleosoma o región sí. De esta manera surge el concepto de denominado "código de histonas", en las que varias modificaciones de histonas actúan de modo secuencial y gradual a la hora de controlar un determinado proceso biológico (Jenuwein and Allis, 2001). Actualmente hay descritos más de 50 "estados de la cromatina" basados en la combinación específica de modificaciones de histonas, para los cuales se sugiere un papel biológico particular (Ernst and Kellis, 2010).

Como las modificaciones de las histonas son dinámicas, se han descrito tanto enzimas que catalizan estas modificaciones como aquellas que las remueven. Las metiltransferasas de histonas (HMTs), desmetilasas de histonas (HDMTs) y las quinasas de histonas actúan sobre residuos y subunidades de histonas específicos. Mientras que la mayoría de las acetiltransferasas (HATs) y las desacetilasas de histonas (HDACs) no son tan específicas y modifican más de un residuo e incluso sobre proteínas distintas de histonas (Kouzarides, 2007).

#### <u>ARN no codificantes</u>

Los ARNnc también juegan un papel fundamental en la regulación epigenética de la expresión génica. Aunque los genomas eucariotas transcriben hasta el 75% del ADN genómico, aproximadamente 3% de estos transcriptos codifican para proteínas; la mayoría son ARNnc, los cuales se clasifican de acuerdo a su tamaño y función (Bernstein et al., 2012). Los ARNnc regulatorios como: ARN pequeños de interferencia (siARN), microARNs (miARN), ARNnc largos (lncARN), actúan en la regulación de la expresión génica a distintos niveles: transcripción, degradación de ARNm, splicing y traducción (Kaikkonen et al., 2011).

Por ejemplo, los miARNs son ARN regulatorios de 20-30 nucleótidos de largo, que hibridan perfectamente con las regiones 3´-no traducidas de los ARNm dianas, resultando en su degradación o inhibición de la traducción del ARNm (He and Hannon, 2004). Se ha estimado que los miARNs regulan la tasa de traducción de más del 60% de los genes codificadores de proteínas, y por lo tanto participan en la regulación de todos los procesos celulares (Friedman et al., 2009).

Como veremos más adelante, estos miARNs están masivamente desregulados en el proceso tumoral. Al igual que los genes codificadores de proteínas, las expresión de las secuencias codificantes de los miARNs pueden ser reguladas por metilación de sus promotores (Lujambio et al., 2008).

#### 1.1.2 Regulación de la expresión génica de manera cooperativa

La metilación del ADN y las modificaciones de histonas son dos mecanismos que interactúan juntos en la regulación de la expresión génica.

La metilación del ADN modula la expresión génica tanto de manera directa como indirecta. El modo más simple implica que la metilación de secuencias de ADN específicas puede inhibir la transcripción impidiendo la unión de algunos factores de transcripción a sus sitios blancos. Un mecanismo alternativo involucra una serie de interacciones ADN-proteína y proteína-proteína, e implica cambios en la arquitectura de los nucleosomas (Figura 5). Así, la metilación del ADN recluta proteínas de unión al ADN metilado (MBD), que preferencialmente se unen a metil-CpGs. Algunas de estas

proteínas, como MeCP2, reconocen el ADN metilado y reclutan varias enzimas modificadoras de histonas a la región. Un grupo importante de estas enzimas son las HDACs, las cuales promueven la condensación de la cromatina e inactivación transcripcional al catalizar la remoción de los grupos acetilos del *core* de histonas. A su vez, las HDACs pueden interactuar directamente con DNMTs (Robertson and Wolffe, 2000).



**Figura 5.** Mecanismo mediante el cual la metilación del ADN y la desacetilación de histonas cooperan para reprimir la transcripción (Robertson and Wolffe, 2000).

# 1.1.3 Reprogramación epigenética durante el desarrollo embrionario y mantenimiento de las marcas epigenéticas

#### Reprogramación epigenética en el embrión temprano

La metilación del ADN sufre cambios dinámicos durante la embriogénesis temprana para inicialmente establecer un estado global desmetilado (borrado de los patrones epigenéticos preestablecidos en los gametos) y luego, progresivamente una ganancia de metilación linaje-específica que mantiene la identidad celular y la estabilidad genómica.

Los patrones de metilación heredados de los gametos son globalmente borrados luego de la fertilización desde las primeras etapas de clivaje hasta el blastocisto, antes de restablecerles luego de la implantación del embrión (Reik et al., 2001). Este proceso ocurre vía múltiples mecanismos: el ADN paterno sufre una rápida pérdida activa de 5mC en el cigoto antes de la primera división celular por conversión de 5mC a 5-hidroximetilcitosina (5hmC) por acción de la hidroxilasa TET3, mientras que el ADN materno es desmetilado de manera pasiva en varias divisiones celulares, probablemente como resultado de la pérdida de la actividad de la DMNTs de mantenimiento (Guibert and Weber, 2013; Wossidlo et al., 2011). Esta reprogramación de la metilación del ADN culmina con un genoma globalmente desmetilado en la masa celular interna (ICM) en el embrión, y se correlaciona con el establecimiento de la pluripotencialidad de las células. Así, la reducción de los niveles genómicos de 5mC en el embrión pre-implantación genera un estado epigenético propicio para la subsecuente especificación del linaje mediante metilación *de novo* progresiva hasta establecer la identidad celular (Hackett and Surani, 2013).

Luego de la implantación del embrión, la metilación del ADN se restablece progresivamente en regiones pobres en CpG y en un pequeño grupo de islas CpG por acción de la DNMT3A y DNMT3B. Sin embargo, hay secuencias que escapan a la desmetilación luego de la fertilización, en particular, loci improntados y secuencias repetidas (Guibert and Weber, 2013).

Mientras la mayoría de los estudios sobre los patrones de metilación durante el desarrollo se han enfocado fundamentalmente en promotores, es probable que la dinámica de 5mC en regiones intragénicas e intergénicas y elementos repetidos también

contribuyan a la restricción del linaje y los fenotipos observados en sistemas deficientes de DNMTs. De hecho, solo un número relativamente pequeño de promotores, principalmente de densidad CpG intermedia, exhiben metilación *de novo* linaje-dependiente durante el desarrollo, sugiriendo que la metilación del ADN en otros contextos genómicos también es relevante durante el desarrollo embrionario (Hackett and Surani, 2013).

Una excepción al proceso anteriormente descrito ocurre en las células germinales primordiales (PGCs) del embrión donde el proceso de reprogramación epigenética es diferente. Dado que las PGCs derivan de células somáticas deben borrar los perfiles de metilación, cambiar las modificaciones de histonas e intercambiar variantes de histonas. Esto incluye el borrado de la impronta genómica de origen parental y de la metilación en cuerpos de los genes, elementos móviles así como de promotores que se encuentran metilados en el embrión temprano (Guibert and Weber, 2013). Así, el borrado de la metilación en PGCs es más completo que en el embrión preimplantado. Esta desmetilación global en PGCs es necesaria para restablecer el estado epigenético compatible con la expresión del programa de la línea germinal y restablecer la impronta genómica en los gametos de acuerdo al sexo del embrión. Siguiendo la reprogramación, las PGCs entran en una fase de remetilación global que resulta en gametos maduros que presentan genomas altamente metilados (espermatozoides) y parcialmente metilado (oocito).

#### Desregulación epigenética y enfermedades humanas

Siendo el desarrollo embrionario temprano la etapa de reprogramación epigenética, no cabe duda suponer que este es un período crítico de susceptibilidad a desregulación epigenética. A su vez, luego de establecida la metilación del ADN y otras marcas epigenéticas tanto en la línea somática como en la germinal, estas deben ser mantenidas a lo largo del curso de las futuras divisiones celulares. La desregulación epigenética puede resultar de errores en los mecanismos de control epigenético o de alteraciones en las propias marcas epigenéticas, tanto en genes improntados como no improntados. Mutaciones germinales en las proteínas involucradas en el establecimiento, borrado, o lectura de las marcas epigenéticas pueden afectar negativamente el desarrollo de múltiples órganos o tejidos, como se puede observar en una variedad de síndromes genéticos humanos (Berdasco and Esteller, 2013; Huidobro et al., 2013).

Una alteración epigenética o "epimutación" se refiere a un patrón de metilación o modificación de histona aberrante. Estas alteraciones pueden ocurrir en ausencia o presencia de una mutación o alteración genética y se denominan epimutación primaria o secundaria, respectivamente. Una epimutación primaria puede resultar de un borrado, establecimiento o mantenimiento erróneo de las marcas epigenéticas (revisado en Inbar-Feigenberg et al., 2013). En genes no improntados, las epimutaciones somáticas causadas por errores mitóticos en el normal mantenimiento de las marcas epigenéticas pueden resultar de una regulación anormal del crecimiento celular. Como veremos a continuación, ejemplo de ello es el cáncer, donde la combinación de alteraciones genéticas y epigenéticas conllevan el crecimiento incontrolado de la célula.

# 1.2 Papel de la metilación del ADN en el desarrollo tumoral

El cáncer es la consecuencia de cambios genéticos y epigenéticos combinados, en muchos casos con factores ambientales, que inducen la activación y/o inactivación de genes específicos que llevan a la transformación neoplásica. En tumores, la hipermetilación de genes supresores de tumores es tan frecuente como sus mutaciones. En las células tumorales, cientos de genes están hipermetilados (metilación del ADN) o con modificaciones de histonas aberrantes y la expresión de miARNs desregulada, mientras que aproximadamente 10 genes están afectados por mutaciones (Choi and Lee, 2013). Esto marca la importancia de las alteraciones epigenéticas en la tumorigénesis.

Si bien las alteraciones epigenéticas descritas en los distintos tipos de cáncer involucran todos los mecanismos epigenéticos: metilación del ADN, modificaciones de histonas, ARNnc y remodeladores de la cromatina, a continuación nos centraremos fundamentalmente en el papel que juega la metilación del ADN en el desarrollo del tumor, dado que es el tema en el que se centra el trabajo de tesis.

Las células malignas muestran importantes alteraciones en sus perfiles de metilación del ADN. Durante la carcinogénesis, el genoma simultáneamente procede a una hipermetilación restringida regionalmente de islas CpG, afectando así las regiones promotoras de genes (principalmente en genes de mantenimiento celular), y una hipometilación genómica ampliamente distribuida. Estas alteraciones generan una ventaja selectiva para la célula tumoral incipiente. Estos cambios epigenéticos pueden provocar el silenciamiento de genes supresores de tumores y miARNs por hipermetilación, la pérdida de imprinting y, menos frecuente, la activación de oncogenes por desmetilación, así como el aumento de la inestabilidad genómica y una aumentada tasa de mutación en dinucleótidos CpG metilados (Figura 6) (Ballestar and Esteller, 2008).



**Figura 6.** El cáncer es el resultado de la combinación de alteraciones genéticas y epigenéticas. La hipermetilación de promotores de genes supresores de tumores resulta en una inactivación estable. Este mecanismo también inactiva microARNs, los cuales regulan oncogenes. La hipometilación global ha sido asociada con inestabilidad genómica. Las mutaciones genéticas actúan tanto inactivando genes supresores de tumores como activando oncogenes. Modificado de Ballestar and Esteller, 2008.

#### 1.2.1 Hipermetilación regional

#### Inactivación de genes codificantes de proteínas

Al día de hoy se han identificado muchos genes con hipermetilación aberrante en sus promotores en prácticamente todas las formas de tumores. Algunos de estos genes susceptibles incluyen reguladores del ciclo celular, genes de reparación del ADN, genes asociados a apoptosis, a regulación hormonal, metástasis, y angiogénesis. Mientras que algunos genes como p16 están metilados en muchos cánceres, otros genes están metilados en tipos específicos de tumores (Sadikovic et al., 2008). Con el surgimiento de los dato de secuenciación masiva y varias iniciativas en el campo de la epigenética como el *"Human Epigenome Project"*, la lista de genes alterados por mecanismos epigenéticos aumenta rápidamente (Laird, 2010).

De acuerdo con la hipótesis de Knudson de los 2 hits, la metilación del promotor puede ocurrir en un gen actuando en el alelo *wild-type*, mientras que el otro alelo está mutado, contribuyendo a la inactivación bialélica de los genes supresores de tumores, tanto como hit primario o secundario en las formas familiares y esporádicas de cáncer (Herman and Baylin, 2003). En este caso, los cambios genéticos y epigenéticos colaboran para inhibir la expresión funcional de un gen en las células tumorales. Incluso, en algunos casos ambos alelos se encuentran hipermetilados (Daniel et al., 2011; Herman and Baylin, 2003).

La hipermetilación de promotores se observa frecuentemente en genes supresores de tumores donde también se ha reportado mutaciones en cánceres humanos. Cerca del 50% de los genes que causan formas familiares de cánceres cuando están mutados en la línea germinal, presentan silenciamiento génico asociado a metilación en varias formas esporádicas del mismo. Por ejemplo, los genes BRCA1 y serin/treonin quinasa 11, cuyas mutaciones en la línea germinal causan forma familiar de cáncer de mama y colon respectivamente, están a menudo silenciados epigenéticamente en las formas esporádicas de estos mismos tipos de tumores. Se ha descrito que el 10-15% de las mujeres con la forma no familiar de cáncer de mama poseen tumores en los cuales el gen BRCA1 está hipermetilado (Stefansson and Esteller, 2013). Por otro lado, los genes CDKN2A (p16), CDH1, RARB2 y GSTP1 presentan hipermetilación de sus promotores en una frecuencia similar en cáncer de mama de casos familiares (incluyendo portadores de mutaciones en BRCA1 y BRCA2) y en casos esporádicos. Así, los cambios epigenéticos son necesarios aún en cánceres con alteración de la reparación del ADN (como en los tumores de mama con mutación en BRCA1). Esto resalta la importancia del papel de los cambios epigenéticos en adición a las mutaciones genéticas en la contribución al desarrollo tumoral.

El impacto del silenciamiento epigenético en el desarrollo tumoral depende de la relevancia del gen blanco, pudiendo afectar a un gen supresor de tumor, a un factor de transcripción, a un gen involucrado en la reparación del ADN o a un miARN. Por ejemplo, el silenciamiento de factores de transcripción inducido por hipermetilación del promotor, como *RUNX3* en cáncer de esófago y *GATA-4* y *GATA-5* en cáncer gástrico y colorrectal, conduce a la inactivación de todos sus dianas (Sharma et al., 2010).

#### Hipermetilación de miARNs

El silenciamiento de miARNs puede impactar sobre diversos genes, contribuyendo a la tumorigénesis y al fenotipo maligno (Lopez-Serra and Esteller, 2012). El genoma humano codifica más de 1400 miARNs, los cuales frecuentemente tienen como diana
genes relacionados con las funciones biológicas muy variadas, incluyendo ciclo celular, desarrollo, diferenciación, apoptosis, entre otras. Con respecto a su relación con el cáncer, los miARN han sido categorizado en: oncogénicos, supresores de tumores, o miARN contexto-dependiente. En la tumorigénesis, la pérdida de expresión de miARNs supresores de tumores promueve la expresión de los oncogenes blancos, y en muchos casos este silenciamiento se debe a hipermetilación de sus promotores o alteración de modificaciones de histonas. Por ejemplo, el silenciamiento del mir-124A conduce a la activación de su blanco, el gen oncogénico *CDK6*, resultando en la fosforilación e inactivación del gen supresor de tumores *RB1* (Lopez-Serra and Esteller, 2012).

# "Drivers" y "passengers" epigenéticos

En un gran número de loci, los cambios de metilación del ADN influyen el potencial transcripcional más que el estatus transcripcional actual del gen. Por ejemplo, la hipermetilación del ADN asociada a cáncer frecuentemente actúa sobre genes que ya están silenciados en células no tumorales por trimetilación de H3K27. Así, aunque es cierto que algunos genes supresores de tumores presentan metilación *de novo* y son silenciados en cáncer, la gran mayoría de los eventos de hipermetilación afecta a genes ya silenciados en células normales (Kulis et al., 2013; Sproul and Meehan, 2013). Siguiendo la nomenclatura adoptada por los estudios de mutaciones en genomas de cáncer, aquellos genes silenciados *de novo* en el tumor se les denomina "*drivers*" epigenéticos, mientras que los segundos son denominados "*passengers*".

#### Mecanismos de hipermetilación anormal en cáncer

Aunque los mecanismos responsables de la hipermetilación anormal observada en el cáncer no están del todo claro, de los estudios genómicos de tejidos normales y tumorales, surgen dos mecanismos posibles: un proceso activo mediado por la unión de factores específicos a las islas CpG, o un proceso pasivo resultante de la pérdida de protección contra la metilación *de novo* (Sproul and Meehan, 2013).

Una hipótesis es que la hipermetilación en cáncer resulta de la sobre-expresión o aumento de la actividad de las DNMTs. Un análisis reciente reporta que la hipermetilación de algunos genes se correlaciona con el incremento de los niveles de DNTM3B en tumor colorrectal (Steine et al., 2011). Sin embargo, por otra parte se han detectado mutaciones en *DNMT3A* en leucemia mieloide aguda y otros cánceres hematológicos que reducen la actividad enzimática de DNMT3A (Walter et al., 2011). Otra hipótesis es que las histonas y otras marcas asociadas podrían jugar un papel en proteger las islas CpG de hipermetilación. Por ejemplo, en células normales la histona H3K4 trimetilada se anti-correlaciona con la metilación del ADN (Sproul and Meehan, 2013).

Por otro lado, los análisis de los metilomas tumorales han asociado también la desmetilación del ADN pasiva con la hipermetilación anormal de islas CpG. Por acción de las enzimas TET (*ten-eleven translocation*), las 5mC son convertidas a 5hmC (Dahl et al., 2011). La reducción de los niveles de 5hmC es frecuentemente observada en tumores, como melanoma, y la función anormal de la enzima TET ha sido asociada a hipermetilación de islas CpG (Lian et al., 2012).

# 1.2.2 Hipometilación global

#### Activación de retrotransposones, oncogenes y miARN

Una paradoja comúnmente observada en cáncer es que a pesar de la hipermetilación regional de genes supresores de tumores, el contenido global de 5mC disminuye en el genoma tumoral. Menos frecuente que la hipermetilación del ADN en regiones codificantes, también ocurre hipometilación gen específica en algunos cánceres, resultando en la activación de oncogenes potenciales (Veeck and Esteller, 2010). Así, la hipometilación aberrante contribuye a la progresión tumoral por activación de oncogenes y retrotransposones, con el subsecuente aumento de la inestabilidad cromosómica.

Entre los tumores sólidos, la hipometilación global del ADN es más evidente en cáncer de mama con un 50% de los casos con reducción del contenido de 5mC comparado con la contrapartida de tejido normal (Veeck and Esteller, 2010).

La hipometilación en tumores se ha observado en secuencias altamente repetidas, medianamente repetidas y de copia única. Sin embargo, acontece más frecuentemente en secuencias repetidas. Estas incluyen secuencias LINEs (*long interspersed nucleotide*  *element-1*), secuencias moderadamente repetidas como el retrovirus endógeno humano K (HERV-K) y satélites centroméricos (Wild and Flanagan, 2010). De manera interesante, la secuencia altamente repetida SINE (*short interspersed nuclear element*), que constituye el 10% del genoma humano, ha sido reportada en menor frecuencia como hipometilada en tumores. Se propone que esto se debe a que existe una fuerte asociación entre los elementos Alu (SINE) y la metilación intragénica, ya que hay un mayor porcentaje de elementos SINEs dentro de los genes que los elementos LINEs o repetidos satélites (Wild and Flanagan, 2010).

La hipometilación de retrotransposones en células tumorales puede causar su activación transcripcional, permitiendo su relocalización e integración en otros sitios del genoma, provocando mutagénesis insercional. A su vez, la alta homología entre los elementos repetidos puede aumentar la posibilidad de eventos de recombinación normalmente inhibidos por metilación del ADN, contribuyendo así a la inestabilidad genómica (Garrick et al., 1998). Específicamente, las secuencias repetitivas en tándem en el centrómero y sus cercanías juegan un papel en el mantenimiento del ADN compactado como heterocromatina en el punto de asociación de las cromátidas hermanas, permitiendo la estabilidad cromosómica. La hipometilación de estas regiones puede llevar a la descondensación de la cromatina y a reordenamientos cromosómicos mediante translocaciones inestables (Eden et al., 2003).

En algunos tumores, como el cáncer de ovario, la hipometilación anormal de ADN satélite está asociada con la progresión tumoral. Sin embargo, en otros casos como el cáncer de mama, está involucrada en el desarrollo temprano del tumor (Veeck and Esteller, 2010). Como se ha mencionado previamente, aunque es un evento relativamente raro, la hipometilación del ADN puede afectar genes individuales y activar oncogenes, promoviendo la disrupción de procesos celulares normales. Esto se observa fundamentalmente en regiones promotoras fueras de islas CpG de genes normalmente reprimidos (Bird, 2002). Por ejemplo, los genes improntados tienen normalmente una expresión monoalélica, sin embargo, la hipometilación de los sitios CpG de sus promotores conduce a una expresión bialélica y ha sido ligado a la carcinogénesis en cáncer colorrectal (fenómeno conocido como pérdida de imprinting) (Virani et al., 2012). Otros ejemplos son los genes de antígenos de melanoma (*MAGE*), genes de diferenciación celular normalmente expresados en las células germinales masculinas; y en los que su expresión es reprimida por hipermetilación del promotor en

30

las células somáticas. En melanoma maligno, cáncer de mama y varios otros tipos de tumores, sin embargo, ocurre una expresión aberrante de los genes *MAGE* asociada a hipometilación de los promotores (van Doorn et al., 2005).

También por hipometilación de sus promotores, los miARN oncogénicos, los cuales tienen como blanco las vías de inhibición del crecimiento celular, se encuentra a menudo desregulados en tumores. Por ejemplo, miR-21 cuya diana es *PTEN*, está sobreexpresado en glioblastoma; miRNA-155 está activado en cáncer de mama, pulmón y varios cánceres hematopoyéticos (Sharma et al., 2010).

# Mecanismos involucrados en la hipometilación global del ADN

Los niveles de metilación global del ADN son el producto de un balance entre los mecanismos que crean y mantienen las marcas y aquellos que las remueven. Un cambio en este balance pude dar lugar a la pérdida de metilación global como la observada en cáncer. Hay dos teorías sobre cómo ocurre la desmetilación en el proceso tumoral, y no se descarta que ocurran ambas en el proceso tumoral. Una involucra una pérdida "pasiva" del mantenimiento de la metilación del ADN a lo largo de las divisiones celulares y múltiples replicaciones del ADN. La alternativa es una pérdida de la metilación del ADN "activa", independiente a la replicación del ADN.

Una disminución de la metilación del ADN se podría dar por una baja actividad de la DNMT1, lo que llevaría a que se perdieran las marcas con las sucesivas divisiones celulares. Sin embargo, tanto el aumento como la disminución de la expresión de DNMT1 han sido reportados en tumores con hipometilación del ADN. También se ha propuesto que la pérdida de la metilación se daría de manera indirecta como resultado de una cascada de alteraciones epigenéticas involucrando cambios en las modificaciones de las histonas (revisado en Wild and Flanagan, 2010). Por último, se plantea la hipótesis de la "actividad desviada de la metiltransferasa". Esta hipótesis sugiere que la hipometilación global del ADN se debe a una alta afinidad de las DNMTs hacia los sitios de daño del ADN, lo que compromete el correcto mantenimiento de la marca epigenética durante los ciclos de replicación del ADN, provocando que el nivel de metilación disminuya en cada división celular (James et al., 2003). Esto implicaría que el daño del ADN precede a la hipometilación del ADN.

Por otro lado, hoy en día se postula que la desmetilación del ADN puede suceder a través de la actividad de las enzimas TET las cuales convierten 5mC en 5hmC (Dahl et al., 2011). En un modelo, las 5hmC no son mantenidas durante la división celular, ya que son interpretadas por la maquinaria de replicación como citosinas no metiladas y así son propagadas como tal dando lugar a desmetilación pasiva. Otro modelo, sin embargo, plantea que las 5hmC son uno de los intermediarios catalizados por las enzimas TET y que estos intermediarios pueden ser reemplazados con citosinas no metiladas sin la necesidad de división celular a través de la actividad de los mecanismos de reparación del ADN conocidos como reparación de escisión de base (proceso de desmetilación activa) (Stefansson and Esteller, 2013).

# 1.2.3 Inactivación mutacional de genes

Finalmente, otro mecanismos mediante el cual la metilación del ADN aberrante puede contribuir a la transformación maligna está relacionada con el hecho que la citosinas metiladas son más propensas a convertirse en timinas que las citosinas no metiladas (Duncan and Miller, 1980), pudiendo provocar inactivación mutacional de genes. Así, la metilación intragénica en células tumorales puede aumentar la tasa de mutaciones no sinónimas, debido a la deaminación espontánea de las 5mC en los dinucleótidos CpG, la cual causa la transición de C→T. Este fenómeno explica la alta incidencia de mutaciones de transición CpG a TpG observada en genes como el gen supresor de tumores p53 y el receptor LDL humano (Rideout et al., 1990). El análisis de metilación del ADN en las regiones codificantes de BRCA1, RB1 y NF1 mostró una metilación de CpG prevalente, especialmente de los CpG ubicados en los "hotspots" mutacionales de esos genes (revisado en Sadikovic et al., 2008). También, el silenciamiento epigenético de la enzima de reparación del ADN O<sup>6</sup>-metilguanina DNA metiltransferasa (0<sup>6</sup>-MGMT) es otro ejemplo de cómo la metilación anormal puede aumentar las tasas de mutación. La proteína O<sup>6</sup>-MGMT remueve los aductos metilguanina inducidos por carcinógenos en el ADN, los cuales producen mutaciones de transición G $\rightarrow$ A si quedan sin reparar. Los tumores con los alelos  $0^6$ -MGMT silenciados están predispuestos a acumular mutaciones génicas (Esteller et al., 2000).

# 1.3 Efecto de los factores ambientales sobre las marcas epigenéticas

Las modificaciones epigenéticas del genoma proveen un mecanismo que permite la propagación estable de los estados de actividad génica de una generación de células a la siguiente. Debido a que los estados epigenéticos son reversibles, estos pueden ser modificados por factores ambientales, estilos de vida o estímulos exógenos, contribuyendo al desarrollo de fenotipos anormales (Berdasco and Esteller, 2010).

Estudios epidemiológicos y experimentales han demostrado la implicancia de varios factores ambientales y hábitos alimenticios en el desarrollo de una gran variedad de neoplasias (Peto, 2001). Los factores ambientales juegan un papel fundamental en la etiología del cáncer humano incluyendo carcinógenos químicos (como los encontrados en el humo de cigarrillo), contaminantes en la dieta (como aflatoxina B1), y carcinógenos físicos (radiación UV). Algunos hábitos de vida como fumar, consumir alcohol, exceso de exposición al sol, consumo de grasas y stress podrían también contribuir al desarrollo tumoral (Herceg and Vaissière, 2011; Peto, 2001).

Estudios epidemiológicos sugieren una asociación inversa entre la ingesta de folato y el riesgo de cáncer de pulmón, orofaríngeo, de estómago, colorectal, de páncreas, de ovario, próstata y mama (Dolinoy et al., 2007). Debido a que los donadores de grupos metilos derivados de la dieta son necesarios para la síntesis de SAM, que provee los grupos metilos requeridos para la metilación del ADN, los factores ambientales que alteran la nutrición temprana o durante largos períodos de la vida pueden potencialmente influir en el fenotipo adulto mediante alteraciones en la metilación de CpG en regiones genómicas lábiles epigenéticamente. Por ejemplo, dietas deficientes en folato y metionina, los cuales son necesarios para la normal biosíntesis de SAM, conducen a una hipometilación del ADN y pérdida del *imprinting* en *IGF-2* (Waterland et al., 2006). Incluso, variantes genéticas particulares en las enzimas involucradas en el metabolismo del folato y la metionina están asociadas con diferentes niveles de metilación del ADN (Paz et al., 2002). Los niveles de SAM pueden ser modulados por la ingesta de ácido fólico, vitamina B12, así como por el consumo de alcohol.

Un hábito de vida que afecta los patrones de metilación y la progresión del desarrollo tumoral es el humo de cigarrillo, que estimula la hipometilación de genes metastásicos en células de cáncer de pulmón mediante la inhibición de la expresión de la DNMT3b (Liu et al., 2007).

Por otro lado, la exposición a metales, como el arsénico, cadmio, níquel y cromo, está ligada a cambios en la expresión de genes controlado epigenéticamente mediante la interacción con acetiltransferasas y desacetilasas de histonas. Entre los estímulos exógenos, se describió la asociación entre patrones de metilación del ADN alterados con la infección con *Helicobacter pylori* de células epiteliales gástricas, contribuyendo así al riesgo a cáncer gástrico. Además, la inflamación crónica se propuso como posible inductor de metilación aberrante en cáncer de hígado o colon. Sin embargo, el papel causal de estos inductores en los perfiles de metilación del ADN aún no son claros (revisado en Berdasco and Esteller, 2010; Terry et al., 2011). Con respecto a esto, los estudios realizados con individuos genéticamente idénticos son un excelente método para entender la interacción medioambiente-epigenoma.

# 1.4 Epidemiología epigenética

Las marcas epigenéticas, más que cambios heredables en la expresión génica, implican cambios heredables en el potencial de expresión de un gen (Jaenisch and Bird, 2003). Esta distinción es crítica; la expresión génica de una célula no es aislada sino que responde a varias señales extracelulares. Así los mecanismos epigenéticos determinan no solo la expresión génica constitutiva sino también el potencial de alterar apropiadamente la expresión génica en respuesta a estímulos externos.

El campo de estudio de la "epidemiología epigenética", término propuesto por Feinberg y su grupo de investigación (Bjornsson et al., 2004) para describir las variaciones epigenéticas dentro y entre las poblaciones y caracterizar la correlación de factores epigenéticos con el riesgo a enfermedades, está en rápido crecimiento y se ha vuelto más relevante en el estudio de enfermedades complejas. En los últimos años, hubo un desarrollo tecnológico masivo en los estudios epigenéticos que facilitó el avance en esta área, como los análisis genómicos de asociación (EWAS), plataformas de microarreglos de metilación del ADN y secuenciación masiva. El conocimiento en este campo está permitiendo a los investigadores descubrir modificaciones epigenéticas que podrían servir de biomarcadores de enfermedad, mejorando el diagnóstico, pronóstico y estratificación de población de riesgo.

Debido a las interacciones genética-epigenéticas, la epidemiología epigenética no puede ser completamente separada de la epidemiología genética. Esta última se focaliza en el rol de los genes heredados en la etiología de la enfermedad; excepto las mutaciones *de novo*, toda la variación genética es heredada. En cambio, la variación epigenética tiene varias fuentes: *i*) herencia genética y epigenética; *ii*) eventos estocásticos durante el desarrollo embrionario; *iii*) efectos de factores ambientales durante el desarrollo temprano y durante la vida del individuo; y *iv*) el envejecimiento (Waterland and Michels, 2007).

Como veremos en detalle más adelante y se esquematiza en la Figura 7, el proceso de desarrollo está gobernado por factores intrínsecos, pero factores ambientales afectan también los patrones epigenéticos. La combinación de diferentes modificaciones epigenéticas se denomina comúnmente epigenotipo. El epigenotipo influye sobre el fenotipo al nacer. Estas modificaciones epigenéticas deben ser mantenidas durante todos

los ciclos celulares, con el fin de no alterar el epigenotipo(s). Factores intrínsecos juegan un papel importante. Sin embargo, factores ambientales y estilos de vida pueden también tener un impacto en cómo los patrones epigenéticos son mantenidos durante la vida. En caso de efectos ambientales aberrantes, o eventos estocásticos, el epigenotipo puede alterarse. Esto da lugar a una expresión génica alterada y, por ende, en un fenotipo alterado (Figura 7). Así, el fenotipo es determinado por el epigenotipo, el cual puede ser alterado durante el desarrollo embrionario o en la vida postnatal (Feil, 2006).



**Figura 7.** Interacción dinámica entre genotipo, epigenotipo y fenotipo. Los patrones heredables de metilación del ADN y otras modificaciones epigenéticas son establecidos durante el desarrollo, en los diferentes linajes del embrión. Esto está influido por el ambiente y otros factores. El epigenotipo(s) resultante determina la expresión génica, y así el fenotipo. El medioambiente, factores nutricionales y estilo de vida impactan sobre el establecimiento y el mantenimiento somático de los patrones epigenéticos, pudiendo alterar el epigenotipo y así influir sobre el fenotipo (Modificado de Feil, 2006).

Mientras el punto de inicio para los estudios de epidemiología genética es la evidencia de heredabilidad del riesgo a una enfermedad, en la epidemiología epigenética es la evidencia de que la variación epigenética interindividual afecta el riesgo a una enfermedad. Los abordajes de estudio en la epidemiología genética son los estudios de ligamientos en familias que permiten escanear el genoma para asociaciones de haplotipos específicos y la enfermedad, y estudios de asociación poblacionales, en los cuales la variación genética en genes candidatos está relacionada con la enfermedad.

Los estudios basados en familias no tienen tanta utilidad en epidemiología epigenética. La herencia epigenética transgeneracional es, en la mayoría de los casos, un componente minoritario en la variación epigenética interindividual, y aun cuando se detecta agrupamiento familiar del epigenotipo y la variación epigenética (Bjornsson et al., 2008; Chan et al., 2006), es difícil de determinar si es causado por una interacción genotipo-epigenotipo, el ambiente compartido o herencia epigenética.

Los estudios de asociación poblacionales junto con los estudios en gemelos son los más útiles para caracterizar las relaciones entre la variabilidad epigenética y la enfermedad. La discordancia entre pares de gemelos monocigóticos (MZ), los cuales son genéticamente idénticos, permite detectar más fácilmente la etiología epigenética de una enfermedad (Czyz et al., 2012; Waterland and Michels, 2007).

# 1.4.1 Variación en la metilación del ADN y enfermedades complejas

El espectro completo de marcas epigenéticas es aún desconocido, pero potencialmente es enorme considerando que el epigenoma humano diploide contiene más de  $10^8$  citosinas de las cuales aproximadamente  $10^7$  son CpGs, y más de  $10^8$  colas de histonas, que todas pueden variar. En la Figura 8 se muestra los patrones y contextos más comunes en que el ADN metilado puede variar. La variación de metilación del ADN en un sitio CpG único es conocida como posición de metilación variable (MVP), la cual se puede considerar como el equivalente epigenético de un polimorfismo de nucleótido simple (SNP). Con los hallazgos recientes de metilación en sitios no CpG, todas las citosinas podrían ser potenciales MVP. Siendo un evento menos frecuente, los CpGs pueden ser metilados en solo una de las dos hebras del ADN por alelo. Esto se conoce como hemimetilación, y probablemente refleja un retraso en el mantenimiento de la metilación del ADN post-replicación en células en proliferación. Si la metilación es alterada en múltiples sitios CpGs adyacentes, se denomina como región diferencialmente metilada (DMR); los DMRs varían considerablemente en largo, siendo típicamente de < 1 kb pero pueden exceder 1Mb. Por otro lado, encontramos posiciones o regiones que varían en la metilación del ADN dependiendo del origen parental, la presencia de un polimorfismo o como resultado de un evento estocástico (metilación alelo-específica, ASM). Y por último, se distingue la metilación haplotipo-específica (HSM), que corresponden a una región diferencialmente metilada que está definida por un grupo de SNPs co-heredados (un haplotipo) (Rakyan et al., 2012).



**Figura 8.** Diferentes tipos de variación en la metilación del ADN que pueden ser identificados con EWAS (estudios de asociación epigenéticos). Simplificando, los casos y controles se asumen que tienen CpGs en estado metilado o desmetilado. Muestras reales contienen poblaciones de células diferentes y así niveles de metilación más heterogéneos a lo largo de un rango dinámico entre 0-100%. MVP: posición de metilación variable; DMR: región diferencialmente metilada; ASM: metilación alelo-específica; HSM: metilación haplotipo-específica (Rakyan et al., 2012).

El papel de la variación de la metilación del ADN en enfermedades complejas ha sido fundamentalmente estudiado en el contexto de cáncer, ya desde los primeros estudios EWAS (Figura 8). Esta variación epigenética puede ser causal de la enfermedad o consecuencia de la misma. Un punto importante es determinar si la variación está presente antes de cualquier signo de la enfermedad. En este contexto, es útil considerar como la variación epigenética puede surgir antes de la enfermedad. Primero, puede ser heredada y así estar presente en todos los tejidos incluyendo la línea germinal (herencia epigenética transgeneracional). Segundo, puede surgir por azar y estar presente en todo el soma si ocurre temprano en el desarrollo (*in útero*), o limitado a uno o pocos tejidos si ocurre post-nacimiento o durante la vida adulta (Czyz et al., 2012; Fraga et al., 2005). Tercero, puede ser inducida por el medio ambiente y estilo de vida durante la vida del individuo o incluso influencias ambientales *in útero* (Christensen et al., 2009). Es posible también que determinados genotipos influyan sobre la variación epigenética (Gibbs et al., 2010; Zhang et al., 2010). Finalmente,

también hay que tener en cuenta la posibilidad de que en algunos casos la variación epigenética asociada a la enfermedad puede ocurrir antes de la manifestación de la enfermedad, pero aún no ser causante de la enfermedad *per se*. Este tipo de fenómeno puede ser debido a confusores, cuando un factor ambiental como fumar, o una variante genética, induce tanto el estado epigenético aberrante como la enfermedad (Rakyan et al., 2012).

# 1.4.2 Herencia epigenética transgeneracional

Aunque los estados epigenéticos normalmente son mantenidos durante la vida del organismo, es raro que estas marcas epigenéticas pasen a la siguiente generación. La transmisión de cualquier información tanto genética como epigenética entre generaciones de individuos requiere la transmisión a través de la línea germinal y alteraciones permanentes en la secuencia de ADN o el epigenoma, respectivamente. Debido a la reprogramación de la metilación del ADN en la fertilización, los sitios epigenéticos modificados necesitarán ser del tipo improntados para escapar del proceso de desmetilación (Skinner, 2011). Los fenotipos heredados epigenéticamente requieren la pérdida de la exposición directa a factores ambientales para ser considerada transgeneracional.

La herencia transgeneracional de la metilación del ADN ha sido descrita en plantas, levaduras, Drosophila, y modelos murinos tanto para transgenes como alelos endógenos (revisado en Youngson and Whitelaw, 2008). Sin embargo, la herencia de metilación del ADN en humanos ha sido muy poco evidenciada y fundamentalmente evaluada en familias con una historia de cáncer. Las epimutaciones constitucionales en los genes de reparación *MLH1* y *MSH2* en el síndrome de Lynch, o cáncer colorectal no poliposis hereditario (HNPCC), representan ejemplos claros de este fenómeno. Mientras las epimutaciones en *MLH1* típicamente ocurren espontáneamente y muestran un patrón de herencia no mendeliano ya que existe en algunos casos reversión durante la gametogénesis, por el contrario, las epimutaciones en MSH2 ocurren secundariamente debido a un defecto genético cercano en *cis* y así presenta una herencia mendeliana. Estas dos formas de epimutaciones representan las dos formas más estudiadas en enfermedades humanas (Hitchins, 2010). Estudios de familias con HNPCC han

mostrado hipermetilación de la región promotora de los genes *MSH2* y *MLH1* en el caso índice y sus hijos afectados (Hitchins et al., 2007).

#### 1.4.3 Variación epigenética durante el desarrollo

Estudios epidemiológicos humanos y datos de modelos animales sugieren que el desarrollo temprano afecta el riesgo de desarrollo de enfermedades complejas en la edad adulta como diabetes, enfermedades cardiovasculares, algunas formas de cáncer, o alergias, entre otras (Hanson et al., 2011), lo que llevó a postular la "Teoría del origen temprano de las enfermedades". Durante los períodos críticos del desarrollo prenatal y posnatal de mamíferos, la nutrición y otros estímulos ambientales influyen las vías normales de desarrollo e inducen así cambios permanentes en el metabolismo y susceptibilidad a enfermedades crónicas.

Los primeros datos empíricos provienen de experimentos en modelos animales, los ratones agouti amarillos (Dolinoy, 2008; Waterland and Jirtle, 2003). Esta cepa agouti surgió por la inserción de un elemento móvil, el retrotransposón intracisternal A, corriente arriba del gen agouti. El gen agouti salvaje codifica una molécula de señalización que produce pigmentación amarilla en el pelaje. La presencia del retrotransposón transforma a la región genómica sensible a la metilación, aunque la razón de esto no es clara. Cuando ratones agouti portadores del elemento móvil son expuestos prenatalmente a altos niveles de donadores de grupo metilo mediante la alimentación de sus madres con dietas ricas en ácido fólico, vitamina B<sub>12</sub> y betaína durante la gestación, esta condición ambiental produce un nivel mayor de metilación del ADN en el locus agouti. Este cambio en la metilación del ADN puede ser observada en muchos tejidos sugiriendo que el cambio epigenético es propagado en todo el cuerpo debido a la transmisión de la metilación del ADN durante la mitosis. Incluso, el cambio epigenético es mantenido de manera estable ya que persiste en la adultez. Esta alteración epigenética en el gen agouti produce cambios en el color del pelaje de los ratones portadores del transposón, y reduce la incidencia de obesidad, diabetes y cáncer. Así, por primera vez, se asoció directamente el efecto de la dieta de la madre durante la gestación sobre el fenotipo adulto de su descendencia con cambios en la metilación del ADN. Esta observación abrió las puertas al estudio del papel de los cambios epigenéticos debido al ambiente durante el desarrollo temprano en la etiología de las enfermedades del adulto (Jirtle and Skinner, 2007).

Las señales recibidas durante el desarrollo embrionario, fetal y en etapa postnatal temprana inducen respuestas adaptativas que permiten el desarrollo de características fenotípicas apropiadas para el ambiente en el cual la descendencia predice que vivirá. Estas respuestas, las cuales se han denominado respuestas adaptativas predictivas (Gluckman et al., 2005), son causadas a través de los mecanismos de la plasticidad del desarrollo, que incluyen procesos epigenéticos. Así, el ambiente produce cambios permanentes en el epigenoma que luego continúan a lo largo del desarrollo y resultan en una modificación del epigenoma y la actividad del genoma (transcriptoma) del adulto. Dado que las respuestas adaptativas son dependientes de la historia de desarrollo y serán diferentes en cada individuo, contribuyen a las diferencias en el riesgo a la enfermedad entre los individuos que aparentemente tienen similar estilo de vida adulta (Hanson et al., 2011).

Aunque las respuestas adaptativas predictivas tienen como objetivo conferir una ventaja adaptativa, los cambios embrionarios que se hacen en base a predicciones en etapas muy tempranas pueden volverse inapropiados para el futuro. Este fenómeno se denomina "development mismatch" (Hanson et al., 2011). Ejemplo de ello es el "Fenotipo Ahorrativo" ("Thrifty Phenotype") propuesto por Hales y colaboradores (1991). Ellos teorizaron que si un feto en crecimiento era mal alimentado durante su gestación debido a una mala nutrición de la madre, disfunción placentaria, estrés u otros factores, el feto se adaptará dentro del útero para aumentar sus chances de sobrevivir si las condiciones de nutrición pobre continúan postnatalmente. El feto maximizará la eficiencia metabolizadora con respecto al almacenaje y uso de nutrientes. Sin embargo, si el ambiente nutricional pre y postnatal no coinciden, esta reprogramación puede ser perjudicial. Varios estudios posteriores han apoyado esta hipótesis, incluyendo estudios de MZ discordantes para la presencia de enfermedades y estudios de niños de madres que vivieron la hambruna danesa durante la Segunda guerra mundial (revisado en Barnes and Ozanne, 2011).

Sin embargo, la primera pregunta que se plantea es si la exposición prenatal, por ejemplo a hambruna, puede dar lugar a cambios persistentes en el epigenoma humano adulto. Individuos que fueron expuestos a hambruna durante la gestación temprana tienen niveles de metilación menor en IGF2 que los controles, seis décadas después de la exposición (Heijmans et al., 2008). Luego se extendió esta observación analizando un grupo de 15 loci candidatos adicionales, implicados en crecimiento, metabolismo y fenotipos cardiovasculares. La metilación en seis de estos loci estaban asociados con la exposición prenatal a la hambruna (Tobi et al., 2009). Aunque las diferencias en la metilación del ADN luego de la hambruna son comunes, el tamaño de las diferencias observadas puede considerarse relativamente pequeño. Este efecto pequeño podría considerarse un fenómeno general ya que también fue observado en estudios epigenéticos humanos de abuso en la niñez temprana (McGowan et al., 2009) y tecnologías de reproducción asistida (Katari et al., 2009), pero los efectos combinados de pequeños cambios epigenéticos podrían resultar en alteraciones en la regulación fina de vías metabólicas, u otras vías implicadas.

Es también posible que no todas las variaciones epigenéticas ocurridas durante el desarrollo embrionario se deban a la influencia del medio externo e interno. Como discutiremos más extensamente en el Capítulo 3, también existe un efecto genético sobre la variación en la metilación del ADN. Los loci que presentan variantes genéticas que influyen el estado de metilación se denominan "*methylation quantitative trait loci*" o methQTLs (Zhang et al., 2010). Un genotipo dado genera una mayor probabilidad de metilación. Feinberg e Irizarry (2010) recientemente reportaron la existencia de variantes genéticas en el genoma humano y murino que no cambian el fenotipo medio sino la variabilidad del fenotipo; y que sería mediado epigenéticamente mediante regiones de metilación variable. La existencia de methQTLs provee un fuerte argumento para integrar GWAS/EWAS para descubrir genotipos que ejerzan su función a través de la variación epigenética.

Al contrario de los cambios que ocurren con la edad, los cambios embrionarios en el epigenoma pueden afectar varios tejidos lo que aumenta las posibilidades de estudios del papel de estos cambios en enfermedades humanas. Para algunos loci, particularmente aquellos cuyo estado epigenético es independiente de la diferenciación celular, tejidos periféricos (como la sangre) pueden ser utilizados para estimar las consecuencias de los cambios epigenéticos ocurridos durante en el desarrollo involucrados en la patología (Ally et al., 2009).

42

# 1.4.4 Variación epigenética durante el envejecimiento

# <u>Deriva epigenética</u>

Los estudios en mellizos monocigóticos demostraron que los patrones de modificaciones epigenéticas de individuos genéticamente idénticos divergen con la edad, es decir con el paso del tiempo los patrones de metilación de los mellizos monocigóticos se diferencian determinando diferencias a nivel fenotípico (Fraga et al., 2005). Estas diferencias en los patrones epigenéticos podrían explicarse por la influencia de factores tanto externos como internos. Así, individuos adultos genéticamente idénticos que han pasado más tiempo separados y han tenido estilos de vida más distintos presentan la mayor diferencia en los patrones de metilación del ADN, probablemente por haber estado sometidos a mayores diferencias del ambiente a lo largo de sus vidas. Es posible que pequeños defectos en la transmisión de la información epigenética a través de sucesivas divisiones celulares o en el mantenimiento de esta información en células diferenciadas se acumulen a lo largo de los años en un proceso que podría ser considerado como una "deriva epigenética" asociada con el proceso de envejecimiento (Fraga et al., 2005; Teschendorff et al., 2013). La acumulación de defectos epigenéticos ocurriría probablemente a una tasa más rápida que la correspondiente a las mutaciones genéticas debido a que sus consecuencias son probablemente menos dramáticas y las células no han desarrollado una cantidad de mecanismos comparables para corregirlos (Fraga et al., 2005). De esta manera, resulta importante analizar si las condiciones ambientales a las que un individuo ha estado expuesto a lo largo de su vida influyen en cambios en los patrones de metilación del ADN y si estas alteraciones epigenéticas predisponen a largo plazo a los individuos a enfermedades como el cáncer. El hallazgo de patrones de metilación diferenciales en el ADN de pacientes con cáncer en comparación con individuos sanos, podría marcar una susceptibilidad a esta enfermedad debido a la acumulación de errores en los eventos epigenéticos normales. Sin embargo, no toda la variación de la metilación del ADN observada durante la vida del individuo se debe a efecto del medioambiente externo e interno. Algunas de las variaciones epigenéticas son estocásticas durante el desarrollo y el envejecimiento. Cambios al azar en la cromatina o en los patrones de metilación durante la vida del individuo o de una generación a la siguiente, pueden contribuir enormemente a la variación total en epigenotipos. El ejemplo mejor comprendido de eventos epigenéticos estocásticos es la inactivación del cromosoma X, en donde la elección del X inactivo es al azar (Bjornsson et al., 2004).

# Pérdida de la plasticidad fenotípica

Un nuevo paradigma en la investigación del envejecimiento surgió recientemente: el envejecimiento podría también derivar de un decaimiento en la capacidad multipotente de células madres adultas. Se han identificado células madres adultas en diversos tejidos, y contribuyen a la homeostasis tisular tanto reparando células dañadas como promoviendo la plasticidad del tejido. Algunos tipos de tejidos que tienen una tasa de recambio alta, como las células sanguíneas o intestinales, presentan gran cantidad de células madres adultas que contribuyen a la constante regeneración de células somáticas. Sin embargo, en tejidos con baja tasa de regeneración, las células madres adultas contribuyen también a la respuesta del organismo a factores ambientales y a la plasticidad tisular (Berdasco and Esteller, 2012). El compromiso de la función de las células madres adultas causa la pérdida de la homeostasis tisular y, consecuentemente pérdida de los fenotipos celulares normales. Nuevas evidencias sostienen que la senescencia de las células madres adultas y el envejecimiento son regulados por modificaciones epigenéticas específicas (Bocker et al., 2011).

Por un lado, se produce una disminución global de la metilación del ADN en diferentes tejidos humanos durante el envejecimiento (Bjornsson et al., 2008). Esta pérdida de metilación está atribuida a una pérdida progresiva en la metilación del ADN en secuencias repetidas, especialmente elementos Alu, localizadas a lo largo de todo el genoma. Sin embargo, esta hipometilación observada en tejidos envejecidos no puede ser explicada por diferencias significativas en la expresión de DNMTs (Maegawa et al., 2010). Paralelamente, y de manera similar a lo que ocurre en el cáncer, ciertos genes son hipermetilados. La hipermetilación relacionada con la edad fue descrita en varios genes regulados en el desarrollo embrionario y algunos de ellos pertenecientes a las vías de senescencia y apoptosis (Bell et al., 2012; Berdasco and Esteller, 2012). Debido a que el envejecimiento es uno de los factores de riesgo más importantes para el cáncer, una predisposición relacionada con la edad para alteraciones en los patrones de metilación del ADN podría ser uno de los factores que aumentarían el riesgo de desarrollo de eventos malignos en algunos individuos.

# 1.5 Biomarcadores epigenéticos en cáncer de mama y melanoma cutáneo

Los biomarcadores de cáncer pueden ser divididos en: i) marcadores diagnósticos, para el tamizaje inicial, ii) marcadores pronósticos, los cuales pueden ser usados una vez que el cáncer ha sido diagnosticado, y iii) marcadores predictivos, los cuales predicen la probabilidad de respuesta a un tratamiento. Por otro lado, se encuentran los biomarcadores de riesgo, los cuales permiten estratificar la población en grupos de bajo y alto riesgo para el desarrollo de determinado tipo de cáncer. Los biomarcadores implican la utilización de herramientas moleculares para su detección, ya que incluyen proteínas, péptidos, ADN, ARNm, miARN, que permiten medirlos de manera cualitativa o cuantitativa. Estos marcadores pueden ser encontrados en el tejido afectado, células sanas y/o fluidos corporales, dependiendo el caso.

Como veremos a lo largo de este trabajo, los biomarcadores epigenéticos han demostrado tener potencial para la evaluación de riesgos, la prevención, la detección precoz, diagnóstico, respuesta al tratamiento, pronóstico y recurrencia en la epidemiología del cáncer. Estudios epidemiológicos basados en biomarcadores epigenéticos, comportamiento y estilo de vida son cruciales para predecir e identificar poblaciones de alto riesgo al desarrollo de cada tipo de cáncer.

# 1.5.1 Cáncer de mama

# Epidemiología y panorama actual en Uruguay

El cáncer de mama es la causa más común de muertes relacionadas a cáncer en mujeres en la mayoría de los países de América. Específicamente, en nuestro país el cáncer constituye la tercera causa de muerte general luego de los accidentes de tránsito y las patologías cardiovasculares; y el cáncer de mama es el tipo de cáncer más frecuente en mujeres uruguayas. De acuerdo a los datos más recientes, en Uruguay la tasa de mortalidad anual estandarizada por edad del cáncer de mama alcanza 22/100.000 mujeres (Luciani et al., 2013) y la tasa de incidencia nacional es de 90,7/100.000

mujeres por año (Globocan 2012, www.globocan.iarc.fr). Estas tasas son las más altas en América Latina y se asemejan a las observadas en los países occidentales desarrollados (Ferlay et al., 2010). Las razones para esta alta frecuencia de cáncer de mama en Uruguay aún no han sido muy bien establecidas. Sin embargo, un número de factores de riesgo ocurren frecuentemente en la población uruguaya, algunos de los cuales han sido implicados directamente en la epidemiología del cáncer de mama, mientras que otros su asociación aún no ha sido totalmente establecida. Dentro de los factores de riesgo asociados a este tipo de cáncer se incluyen la etnicidad, la edad avanzada, historia familiar de cáncer de mama, menarca temprana, menopausia tardía, alimentación, obesidad, y mutaciones en genes de alta penetrancia como *BRCA1*, *BRCA2* y p53, entre otros (Dumitrescu and Cotarla, 2005).

En Uruguay, los estudios de cáncer de mama conducidos hasta hoy en día se han focalizado en identificar factores ambientales y hábitos de riesgo (dieta, actividad física, consumo de alcohol, entre otros) (Ronco et al., 2010, 2011, 2012; De Stefani et al., 2012) así como mutaciones altamente penetrantes en familias afectadas (Delgado et al., 2002, 2011). Sin embargo, las mutaciones en genes de susceptibilidad al cáncer de mama, *BRCA1 y BRCA2* y, menos frecuente *CHEK2, TP53* y *ATM*, se observan solo en el 25% de los tumores de mama familiar (Smith et al., 2006). El restante 75% de casos con cáncer de mama familiar y la mayoría de los tumores esporádicos no son atribuibles a mutaciones conocidas en ninguno de estos genes.

# <u>Alteración de la metilación del ADN en cáncer de mama</u>

La importancia de los cambios epigenéticos en el desarrollo del cáncer de mama está actualmente bien establecida. Las alteraciones en la metilación del ADN han sido extensamente descritas en tejidos de mama tumoral, más recientemente mediante análisis genómicos amplios que revelaron loci hipermetilados en comparación con tejidos adyacentes sanos (Huang et al., 2011). Se han reportado en la literatura más de 100 genes candidatos con promotores hipermetilados con distintas frecuencias en cáncer de mama, incluyendo *BRCA1, CDKN2A, RASSF1A, ERα, PR, RARβ, CCND2* y *PITX2* (Pubmeth, http://matrix.ugent.be/pubmeth/search.html).

Por otro lado, la hipometilación genómica global en el tumor de mama se ha correlacionado con algunas características clínicas como etapa de la enfermedad,

tamaño del tumor y grado histológico (Soares et al., 1999). Algunos protooncogenes implicados en la proliferación y metástasis o en la resistencia a la terapia hormonal se encuentran sobreexpresados debido a la hipometilación de sus promotores (Jovanovic et al., 2010; Lo and Sukumar, 2008).

A pesar de que en la última década se han realizado progresos significativos en la identificación y caracterización de los patrones de metilación del ADN alterados en el desarrollo del cáncer de mama y su progresión, el papel de los cambios epigenéticos como mecanismo de aumento del riesgo de cáncer aún no es claro. Determinar los patrones de metilación en el genoma de las células mamarias normales y tumorales, así como en células sanguíneas normales es fundamental para el futuro desarrollo de biomarcadores de detección temprana y clasificación de los subtipos de cáncer de mama y pronóstico, así como indicadores de riesgo en la población.

Si bien existen algunos datos de posibles marcadores epigenéticos asociados a riesgo a cáncer de mama, aún son limitados y fueron detectados en otras poblaciones. Dada la alta incidencia de este tipo de cáncer en Uruguay, es fundamental detectar marcadores propios de nuestra población, la cual tiene un contexto genético y ambiental particular.

# 1.5.2 Melanoma cutáneo maligno

#### Epidemiología y panorama actual en Uruguay

El melanoma es un tipo de cáncer único, altamente agresivo, el cual aumenta su frecuencia con la edad, y tiene una contribución significativa de factores ambientales en su etiología. Como uno de los cánceres humanos más agresivos, el melanoma es capaz de dar metástasis letales cuando el volumen del tumor primario es de apenas 1 mm<sup>3</sup> (Garbe and Leiter, 2009). A pesar de representar solo el 3% de todas las neoplasias de piel, melanoma cutáneo es responsable del 65% de las muertes relacionadas con cánceres de piel, y la sobrevida a 5 años de los pacientes con metástasis es de 7-19% (Sigalotti et al., 2010).

El melanoma es una de las neoplasias que muestra una tasa de incidencia en constante crecimiento a nivel mundial, de aproximadamente un 3 a 7% anual (Garbe

and Leiter, 2009). Esto también se observa en Uruguay, donde hasta el año 1996 la tasa de incidencia era de 2,4 por 100.000 hombres y 2,1 en mujeres para aumentar hasta 4,1 y 3,3 respectivamente durante el período 2002-2005 (Comisión Honoraria de Lucha Contra el Cáncer, www.urucan.org.uy) (Barrios et al., 2010)

Al igual que en otros tipos de cáncer, el desarrollo de melanoma depende de una interacción entre la estructura del genoma de una persona y factores medio ambientales y/o comportamentales de riesgo. Principalmente, una exposición prolongada al sol sin protección en forma intermitente, antecedentes de 3 o más quemaduras severas antes de los 20 años, antecedentes familiares de 1<sup>er</sup> o 2° grado, cabello pelirrojo o presencia de lunares entre otros son considerados como los factores de riesgo más importantes. En particular, la exposición solar prolongada e intermitente es señalada como una de las principales causas asociadas al aumento de melanoma en el mundo.

Alrededor del 10% de los melanomas cutáneos ocurren de manera familiar con 2 o más parientes cercanos afectados, indicando la baja prevalencia pero con alta penetrancia de los genes involucrados (Gandini et al., 2005). En el melanoma de tipo familiar, se detectan mutaciones en los genes *CCND1*, *CDKN2A*, *CDK4*, *PTEN*, *MITF* y *MC1R* que explican el desarrollo de la enfermedad, pero solo en el 40% de los casos aproximadamente. Como ejemplo, mutaciones en *CDKN2A* se detectan en < 5% de las familias en que al menos dos individuos, parientes de primer grado, presenten melanoma. Esto aumenta hasta el 40% en los casos donde en la familia se detectan al menos 3 individuos parientes en primer grado afectados. Mientras que 1% corresponde a los genes *CCDN1*, *CDK4*, *PTEN*, *MITF* y *MC1R* (High and Robinson, 2007). En el 59% restante aún no han sido identificados los genes involucrados.

Los alelos de riesgo bajo o moderado con alta prevalencia y baja penetrancia son frecuentes en los casos esporádicos de melanoma (Eggermont et al., 2013). Muchas variantes mutacionales del gen *MC1R*, que influyen el fototipo, están asociadas con un aumento de riesgo de melanoma, dependiendo del SNP (Williams et al., 2011).

En nuestro país pocos estudios se han reportado sobre los mecanismos involucrados en el desarrollo de melanoma cutáneo y su base genética, y la mayoría se han centrado en determinar las frecuencias de mutaciones en genes previamente reportados como asociados a melanoma. Así se determinó que en la mayoría de las familias uruguayas afectadas por melanoma cutáneo, el gen mutado es *CDKN2A* (Larre Borges et al., 2009). Otro estudio uruguayo reportó una alta frecuencia de mutaciones en el gen *BRAF*  en individuos con melanoma cutáneo (Mazzei et al., 2013). Con un abordaje más amplio, nuestro grupo de investigación está trabajando en la identificación de regiones genómicas asociadas a melanoma esporádico, mediante mapeo por mestizaje (Hochmann, 2012; Hochmann et al., 2010).

#### Avances en estudios epigenéticos en melanoma

Recientemente a nivel mundial, ha habido un gran énfasis en identificar y analizar los eventos epigenéticos involucrados en la progresión y metástasis de melanoma. Al día de hoy, se han identificado al menos 50 genes que se encuentran específicamente silenciados por hipermetilación de la región promotora en melanoma comparado con melanocitos normales (Sigalotti et al., 2010). Muchos de los genes que se encontraron metilados en melanoma raramente son blancos de mutaciones (Dahl and Guldberg, 2007). De manera interesante, se ha observado una incidencia muy alta de metilación en promotores de genes involucrados en la activación metabólica de drogas quimioterapéuticas, lo cual contribuye a la resistencia de las células de melanoma cutáneo a la quimioterapia convencional (Muthusamy et al., 2006).

Por otro lado, también se detectó hipometilación global del ADN en muestras de melanoma. Sin embargo, si el grado de hipometilación es suficiente para distinguir los nevos benignos de melanoma aún no es claro (Lian et al., 2012). La hipermetilación gen específica es un mejor discriminador como indican estudios donde perfiles de metilación del ADN multi-locus pueden diferenciar melanomas de nevos (Tellez et al., 2009). Más recientemente, se ha identificado la pérdida de 5-hmC como un evento epigenético clave en melanoma, con implicancias en el diagnóstico y pronóstico, debido a la baja expresión de las enzimas isocitrato deshidrogenasa 2 (IDH2) y familia TET (Lian et al., 2012).

Entre los tipos de cáncer de piel, sin duda que melanoma es el más discutido desde el punto de vista epigenético debido a su agresividad y su probada resistencia a las terapias convencionales. Se han desarrollado varias patentes y ensayos clínicos para el uso de drogas epigenéticas en el tratamiento de melanoma. Sin embargo, a diferencia de otros tipos de cáncer, prácticamente no existen estudios de búsqueda de marcadores de riesgo de base epigenética asociados a melanoma cutáneo.

49

# 1.5.3 Aportes del trabajo de tesis en la búsqueda de biomarcadores de riesgo en cáncer de mama y melanoma cutáneo

En este trabajo de tesis queremos demostrar la existencia de metilación del ADN diferencial en células sanguíneas de pacientes con cáncer. Esto sugiere que la variación en la metilación del ADN observada en cáncer es un mecanismo sistémico que, aun cuando el tejido afectado presente perfiles de metilación particulares necesarios para el desarrollo del fenotipo tumoral, existen marcas epigenéticas en otros tejidos no afectados (en nuestro caso, células sanguíneas) que podrían ser indicadoras de la patología. Estas marcas, aunque no estén relacionadas directamente con el proceso tumoral y expliquen totalmente el fenotipo, podrían aportar información para la identificación de posibles causas y podrían utilizarse como biomarcadores de riesgo. Con el fin de detectar marcadores de susceptibilidad para el desarrollo de melanoma cutáneo y cáncer de mama esporádico en la población uruguaya, integramos la información genética y epigenética de los individuos.

# Hipótesis, objetivos y estrategia general de trabajo

# 2.1 Hipótesis de trabajo

En base a lo expuesto en el capítulo anterior, nuestra hipótesis de trabajo plantea la existencia de cambios epigenéticos en los individuos con cáncer, heredados o adquiridos en etapas tempranas del desarrollo o a lo largo de sus vidas, que dejan una impronta en el ADN de células que no están necesariamente relacionadas con el tumor primario. El estudio de epimutaciones en los tejidos y células que no están afectadas por el proceso de la enfermedad permitiría determinar una asociación causal. Estos cambios epigenéticos sistémicos sucederían en regiones susceptibles que asociados a la ocurrencia de otras alteraciones genéticas y/o epigenéticas en determinado tejido conllevarían al desarrollo tumoral, por lo que pueden ser asociados a una predisposición al cáncer.

# 2.2 Objetivos

# 2.2.1 Objetivo general

Evaluar la existencia de metilación diferencial del ADN en leucocitos de pacientes con cáncer de mama y melanoma esporádico y su localización en el genoma, como posibles biomarcadores de riesgo en la población uruguaya.

# 2.2.2 Objetivos específicos

- Evaluar el nivel de metilación global del ADN genómico en leucocitos de pacientes con cáncer de mama y pacientes con melanoma en comparación a un grupo control no afectado.
- 2. Identificar secuencias específicas diferencialmente metiladas en leucocitos de pacientes con cáncer de mama y pacientes con melanoma.
- Analizar el efecto de características epidemiológicas, genéticas y clínicas de los individuos sobre los niveles de metilación global y niveles de metilación loci-específicas.
- 4. Caracterizar los sitios loci-específicos con metilación diferencial en los pacientes con cáncer, y evaluar dichos sitios mediante meta-análisis utilizando bases de datos públicas de expresión génica.
- 5. Validar los perfiles de metilación de los sitios diferencialmente metilados detectados en pacientes con cáncer en tejidos normales, tejidos tumorales y líneas celulares en una cohorte independiente correspondiente a población europea.
- 6. Seleccionar un grupo de sitios diferencialmente metilados en pacientes con cáncer como posibles candidatos de biomarcadores de riesgo.

# 2.3 Estrategia a seguir

# 2.3.1 Abordajes de análisis

La estrategia de trabajo fue realizar un estudio de asociación entre patrones globales de metilación del ADN y dos tipos de cánceres esporádicos en la población uruguaya, integrando el análisis epigenético y genético. Específicamente el trabajo se centra en cáncer de mama y melanoma cutáneo maligno. Se realizaron dos abordajes distintos que comprenden dos grandes hitos o etapas de este trabajo:  Nivel de estudio Global. Detección de metilación global del ADN mediante cuantificación relativa de citosinas metiladas a lo largo de todo el genoma, con el fin de elucidar si era posible detectar metilación diferencial del ADN genómico en células sanguíneas entre pacientes con cáncer esporádico y su respectivo control.

2) Nivel de estudio Genómico y sitio-específico. Una vez que comprobamos la existencia de diferencias en la metilación del ADN genómico a nivel global en pacientes con cáncer, nos centramos en intentar ubicar la localización genómica de dichas diferencias epigenéticas. Para ello, se estudiaron los perfiles obtenidos mediante Amplificación entre sitios metilados (AIMS) en el estudio de melanoma, y los perfiles de microarreglos de metilación para el estudio de cáncer de mama.

# 2.3.2 Población de estudio

#### Estudio en cáncer de mama esporádico

Se seleccionó ADN procedente de muestras de pacientes con cáncer de mama esporádico y controles no afectados (sanos) de 500 individuos que participaron del proyecto de investigación "Genetic ancestry and breast cancer risk in an admixed latin american population from Uruguay" (financiado por Susan Komen Foundation, POP0601362), donde participó nuestro grupo de investigación (Bonilla et al., 2010). Dichas muestras fueron recolectadas en Hospital de Clínicas, Hospital Pereira Rossell, Hospital Militar, Instituto de Oncología y Casa de Galicia en Montevideo, y cuatro hospitales en el interior del país: Asistencial del Sanatorio Cantegril (Maldonado), Hospital Regional de Tacuarembó (Tacuarembó), Cooperativa Médica de Paysandú (Paysandú) y Hospital Regional de Soriano (Soriano).

Todos los casos fueron seleccionados de acuerdo a los siguientes criterios: mujeres con cáncer de mama mayores de 35 años, sin antecedentes personales de cáncer y sin antecedentes familiares de primer grado de cáncer de mama y/o ovario, y sin padres consanguíneos. Todas las pacientes habían sido diagnosticadas con cáncer de mama un

año o menos antes de incluirse en el estudio, y no estaban sometidas recientemente a radio o quimioterapia en el momento de la toma de la muestra.

El grupo control se confeccionó tomando una muestra entre mujeres sanas mayores de 35 años, sin historial familiar de primer y segundo grado de cáncer, con estudio mamográfico normal de menos de 1 año de antigüedad, no emparentados con ningún otro participante del proyecto. Los controles fueron seleccionados en los mismos hospitales y clínicas que los pacientes.

# Estudio en melanoma cutáneo maligno esporádico

Se seleccionaron muestras de pacientes con melanoma esporádico e individuos controles no afectados, de un pool de 220 individuos que participaron del proyecto de investigación "Identificación y mapeo de regiones del genoma asociadas al desarrollo de melanomas esporádicos" (Proyecto CSIC I+D 2008) llevado a cabo por el Dpto. Genética y la Cátedra de Dermatología del Hospital de Clínicas. Tanto las muestras de pacientes como de los controles de melanoma fueron recolectadas en el Hospital de Clínicas "Manuel Quintela". El grupo control se confeccionó tomando una muestra de entre individuos sanos que no tenían una historia familiar de cáncer, ni estuvieran emparentados con las personas de la muestra de pacientes. El grupo de pacientes estaba pareado con el grupo control por edad y sexo.

## <u>Aspectos bioéticos</u>

Todos los datos y muestras fueron tratados en forma totalmente anónima únicamente identificados por un código que no identifica a la persona participante. Todas las personas fueron informadas de los alcances de los estudios y autorizaron la utilización de su muestra y datos para posteriores estudios asociados a la temática del cáncer. Este proyecto, así como los estudios anteriores con las mismas muestras y el consentimiento informado fueron aprobados por el Comité de Ética de la Facultad de Medicina (Exp. Nº 071140-000303-12, resolución del Consejo de Fac. Medicina de fecha 11 de abril de 2012; Exp.Nº 071140-000439-06, resolución del Consejo de Fac. Medicina de fecha 27 de diciembre de 2006, Exp. Nº 071140-000847-07, resolución del Consejo de Fac. Medicina de fecha 7 de noviembre de 2007, y Exp. N°071140-000569-08, resolución del Consejo de Fac. Medicina de fecha 18 de junio de 2008).

## Datos epidemiológicos

Con la aprobación de los participantes de ambos estudios se realizó una encuesta en persona sobre datos clínicos, familiares y epidemiológicos, y la extracción de una muestra de sangre periférica.

Para este estudio contamos con información epidemiológica exhaustiva obtenida del extensivo cuestionario que fue aplicado a todos los participantes, recabándose información sobre antecedentes médicos, historia personal, historia familiar, hábitos de vida (hábitos alimenticios, fumar, actividad física, consumo de alcohol, etc.), factores ambientales, culturales y sociodemográficos desde su niñez hasta la actualidad.

Sólo algunas variables fueron analizadas para el trabajo de esta tesis: edad, sexo, índice de masa corporal (IMC), estatus de fumador y estadío tumoral. El IMC fue calculado como peso medido (kg) dividido por la altura medida (m) al cuadrado. La edad se analizó como variable continua, mientras que los restantes parámetros del estudio se consideraron como variables cualitativas. El IMC se clasifició en 3 rangos: <25; 25-29,9; >30, mientras que el estatus de fumador se clasificó como si/no considerando al individuo como fumador tanto si lo era en el momento de la toma de la muestra como en el pasado. La obtención de todos los datos no fue posible para todos los individuos incluidos en el estudio.

#### <u>Muestras biológicas utilizadas</u>

Todos los ensayos de laboratorio descritos en este trabajo se realizaron con muestras de ADN genómico almacenadas previamente a -20°C. Dichas muestras fueron extraídas a partir de sangre periférica mediante separación previa de células blancas (leucocitos) y extracción del ADN mediante métodos convencionales. Antes de ser almacenadas se midió su concentración a 260 nm de absorbancia por NanoDrop 1000 (ThermoScientific) y se determinó la relación Absorbancia 260nm/280 nm como medida de su pureza.

# Datos de metilación de tejidos y líneas celulares

Los datos de metilación de sitios CpG de tejidos y líneas celulares fueron cedidos gentilmente por el Dr. Manel Esteller y su grupo de investigación. Se utilizaron los datos obtenidos mediante la plataforma Illumina 450k HumanMehtylation BeadChip de la colección de tejidos del Programa de Epigenética y Biología del Cáncer (PEBC) del Instituto de Investigaciones Biomédicas de Bellvitge (Barcelona). De esta manera, contamos con datos de metilación de más de 450.000 sitios CpG en las siguientes muestras: 16 nevos benignos, 37 tumores primarios de melanoma cutáneo maligno, 36 tumores metastásicos de melanoma, 19 líneas celulares de melanoma comerciales, 67 tumores primarios de cáncer de mama, 18 tejidos sanos de mama y 11 líneas celulares de mama y tejido mamario sano de 14 pacientes. Todas las muestras de tejidos pertenecen a individuos europeos. Las líneas celulares analizas se detallan en la Tabla Suplementaria 1.

# **Boom Service Address of the service of the service**

# **3.1 Antecedentes**

Como se ha descrito ampliamente en el Capítulo 1, alteraciones en los patrones de metilación del ADN, tanto a nivel genómico global como loci específico, han sido observadas en prácticamente todos los tipos de cáncer. Las alteraciones descritas en tejido tumoral incluyen: hipermetilación en promotores de genes y una reducción del nivel global de 5mC (Berdasco and Esteller, 2010; Taby and Issa, 2010). Ya que la carcinogénesis induce numerosos cambios genéticos y epigenéticos, el estudio únicamente de células tumorales no puede distinguir si un cambio epigenético es una causa o consecuencia del proceso neoplásico (Baylin and Bestor, 2002).

# 3.1.1 Metilación global del ADN genómico como marcador de riesgo

Aunque los perfiles de metilación del ADN son en general tejido específicos, datos recientes indican que las marcas epigenéticas en leucocitos de sangre periférica (PBL) son marcadores fenotípicos de inestabilidad genómica y marcadores de riesgo candidatos para varios tipos de tumores sólidos como carcinoma de célula escamosa de cabeza y cuello (Hsiung et al., 2007), cáncer de vejiga (Cash et al., 2011; Moore et al., 2008; Wilhelm et al., 2010), cáncer de mama (Choi et al., 2009), de estómago (Gao et al., 2012; Hou et al., 2010), hígado (Guerrero-Preston et al., 2007) y adenoma y cáncer colorrectal (Lim et al., 2008; Pufulete et al., 2003), aún luego de ajustar los resultados por factores de riesgo a cáncer conocidos (Pedersen et al., 2011; Teschendorff et al., 2009; Woo and Kim, 2012; Zhu et al., 2011) (Figura 9). La mayoría de estos trabajos

encontraron un riesgo elevado para cáncer entre los individuos en el cuantil más bajo de metilación global del ADN genómico comparado con aquellos en el cuantil más alto, independientemente del método utilizado para medir la metilación genómica global. Esta hipometilación global detectada en células sanguíneas de los pacientes con cáncer en comparación a un grupo control no afectado, indicaría que los pacientes con cáncer presentarían una disminución sistémica de la metilación del ADN global. Incluso, Friso y colaboradores (2013) también observaron que la metilación global en PBL al momento de la toma de la muestra fue menor en individuos que desarrollaron cáncer en el período posterior de seguimiento de 8 años cuando se lo comparó con aquellos individuos que no habían desarrollado cáncer durante el mismo período. Los resultados del trabajo de Friso apoyan la hipótesis de que la disminución de la metilación global del ADN indica una condición sistémica de susceptibilidad al desarrollo tumoral en lugar de un efecto directo del cáncer (Friso et al., 2013).



**Figura 9.** Gráfico de Forest de asociación entre riesgos a cáncer ea hipometilación global del ADN en leucocitos de sangre periférica. Methyl: Methyl acceptance assay, F: Modelo de efectos fijos, R: Modelo de efectos aleatorios. Las líneas horizontales a través de los boxes representan intervalos de confianza 95%. El centro de los boxes están situados en el punto estimado (OR/RR) (Woo and Kim, 2012).

Así, el nivel de metilación del ADN genómico de leucocitos podría ser un marcador potencial para establecer el estatus sistémico de metilación del genoma abriendo nuevas posibilidades para generar biomarcadores de susceptibilidad así como en la detección temprana de cáncer (Mulero-Navarro and Esteller, 2008). A diferencia del ADN tumoral que requiere para su obtención una biopsia mediante técnicas

agresivas/invasivas para el paciente, el ADN de leucocitos puede ser obtenido en grandes cantidades de manera no invasiva y a bajo costo.

#### 3.1.2 Variabilidad entre las poblaciones

Sin embargo, los resultados de estos estudios epigenéticos-epidemiológicos no pueden ser extrapolados directamente a otras poblaciones, ya que cada población tiene un contexto genético y ambiental particular.

Existen evidencias que varios grupos étnicos difieren en términos de sus patrones de metilación en tejidos sanos como células sanguíneas, piel, colon y próstata (Figueiredo et al., 2009; Heyn et al., 2013a; Kwabi-Addo et al., 2010; Terry et al., 2008; Winnefeld et al., 2012). Por ejemplo, se ha detectado niveles globales de metilación del ADN de PBL más bajo en mujeres afroamericanas sanas de edad media en comparación a mujeres americanas con ascendencia europea (Terry et al., 2008). Resultados similares fueron reportados también por Zhang y colaboradores (2011) quienes detectaron niveles de metilación global significativamente menores en afrodescendientes no hispanos comparado con caucásicos no hispanos. Aún más interesante, el trabajo de Adkins (2011) sugiere que pequeñas diferencias étnicas en la metilación del ADN existen incluso desde el nacimiento en células de la sangre.

Por otro lado, varios estudios han demostrado la existencia de diferencias en los patrones de metilación también en el ADN tumoral de pacientes con diferente ancestralidad (Das et al., 2006; Wang et al., 2012).

Las tasas de incidencia en algunas enfermedades asociadas a alteraciones epigenéticas, como el cáncer, difieren entre los grupos étnicos y estas diferencias pueden ser explicadas por diferentes exposiciones ambientales, estilos de vida, y/o diferencias intrínsecas en las frecuencias de variantes genéticas o epigenéticas. Ejemplo de ello son el cáncer de mama y melanoma maligno cutáneo.

Con respecto al cáncer de mama, su incidencia varía sustancialmente a lo largo de las distintas poblaciones y grupos étnicos. Este hecho ha sido ampliamente estudiado en Estados Unidos (USA). Específicamente, la ancestría genética está asociada con el riesgo a cáncer de mama en mujeres latinas y mexicanas de USA, donde mayor ancestría europea fue asociada con un aumento del riesgo y una mayor ancestría nativo americana fue asociada con un riesgo menor a cáncer de mama (Fejerman et al., 2008, 2010, 2012).

Por otro lado, la incidencia global de melanoma cutáneo maligno ha estado aumentando continuamente durante las últimas cuatro décadas en poblaciones europeas y descendientes de europeos (Garbe and Leiter, 2009). Al igual que el cáncer de mama, la ancestría europea fue asociada con un mayor riesgo a melanoma cutáneo en Brasil (Luiz et al., 2012). La población uruguaya ha sido descrita fundamentalmente como de origen europeo. Sin embargo, análisis de mestizaje genético demostraron que es una población tri-híbrida con una contribución genética también de nativo americanos y africanos (Hidalgo et al., 2005; Sans et al., 1997). Nuestro grupo de investigación previamente ha detectado un exceso de ancestría europea en pacientes con melanoma uruguayos (Hochmann et al, manuscrito en preparación); sin embargo, muy poco se ha publicado sobre melanoma cutáneo en otras poblaciones mestizadas.

Pocos estudios han analizado la asociación entre etnicidad y metilación del ADN en pacientes con cáncer, y todos ellos fueron basados en datos de ancestría por autodenominación de los individuos. A su vez, la mayoría de los trabajos han investigado cambios de la metilación del ADN que ocurren a nivel del tejido en estado sano y tumoral (revisado en Mohammed et al. 2012).

Con el objetivo de detectar marcadores de susceptibilidad para el desarrollo de melanoma cutáneo y cáncer de mama esporádico en la población uruguaya, integramos la información genética y epigenética de los individuos. Como se describe en este capítulo, analizamos la metilación global del ADN en leucocitos de sangre periférica como posible marcador de riesgo de cáncer esporádico y su asociación con la ancestría genética individual y variantes genéticas en genes candidatos de cáncer

# 3.2 Metodología

# 3.2.1 Cuantificación de metilación global del ADN genómico

Los niveles de metilación global del ADN genómico proveniente de leucocitos de sangre periférica fueron evaluados mediante cuantificación de 2'-desoxi-nucleósidos (Berdasco et al., 2009). El DNA fue hidrolizado con Nucleasa P1 y fosfatasa alcalina para producir 2'-desoximononucleósidos, que fueron separados por cromatografía líquida de alta resolución (HPLC) y detectados por UV.

Brevemente, 1 µg de ADN genómico (en 10 µl de agua ultra pura) de cada muestra fue desnaturalizado a 94 °C durante 10 min y rápidamente enfriado en hielo por 5 min. Luego se adicionó 1 µl de ZnSO<sub>4</sub> 10 mM y 1,5 µl de nucleasa P1 (Sigma-Aldrich, 200 U/ml en C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>Na 30 mM) y se incubó a 37 °C durante 16 h. El ADN hidrolizado se incuba con FastAP<sup>TM</sup> Thermosensitive Alkaline Phosphatase (Fermentas) durante 1 h. adicional a 37 °C y 10 min a 75 °C según indicaciones del proveedor, en una reacción de 50 µl de volumen final.

Para la separación de los nucleósidos tras la hidrólisis genómica se utilizó HPLC (Agilent serie 1100 HPLC system). 50 µl de ADN hidrolizado fueron inyectado en una columna dC18 Atlantis (Waters) de fase reversa (2.1 x 20 mm; 5 mm de tamaño de partícula) protegida por una precolumna (2.1 x 20 mm; 5 mm de tamaño de partícula; Agilent) a un flujo constante de 0.150 ml/min. Se utilizaron 2 buffers: ácido fórmico 0.1% en agua (solvente A) y ácido fórmico 0.1% en agua (50%): metanol (50%) (solvente B); en un gradiente isocrático de 95% del solvente A y 5 % del solvente B durante 25 minutos. Se utilizaron como estándares una mezcla a 5mM de los mononucleósidos siguientes: deoxiadenina (dA), deoxitimidina (dT), deoxiguanosina (dG), deoxicitidina (dC), 5-metil-2'-deoxicitidina (5mdC) y deoxiuridina (dU) (Sigma-Aldrich).

La identificación de dC y 5mdC fue obtenida por detección UV a A254 y A280 (Figura 10). La cuantificación de 5mdC y dC se obtiene directamente del área de los picos visualizados en el HPLC. El nivel de metilación global del ADN genómico se expresa como la cantidad relativa de 5mdC en el total de residuos citidinas, es decir que, el porcentaje de metilación global del ADN genómico es:  $[mdC / (mdC + dC)] \times 100$ .



**Figura 10.** Separación de los deoxinucleósidos por HPLC. Cromatogramas obtenidos por detección UV a una absorbancia a 254 nm y 280 nm.

Todas las muestras fueron analizadas en el HPLC por duplicado, y se eliminaron las muestras cuya diferencia en contenido mdC entre los duplicados fue mayor del 3%. Se calculó la media para cada muestra. En cada experimento se tomó el cuidado de analizar muestras del grupo de pacientes y del grupo control, para minimizar el efecto del error experimental entre los grupos. Los resultados dudosos fueron verificados por espectrometría de masas (ionización por electrospray; ESI-MS).

# 3.2.2 Genotipado de SNPs y determinación de ancestralidad individual

Se utilizó los datos de ancestría individual obtenidos a partir de un grupo de 59 marcadores informativos de ancestralidad (MIAs). Estos MIAs son SNPs bialélicos seleccionados de un panel descrito por Fejerman et al. (2010) (Tabla Suplementaria 2). La selección de MIAs se basó en los cálculos de diferencias de frecuencias alélicas entre europeos, africanos y nativo americanos. Los SNPs seleccionados maximizan la información de apareamiento de más de una población ancestral, con una gran diferencia en frecuencia alélica entre las poblaciones ancestrales (>0.5). Los MIAs utilizados están distribuidos a lo largo de todo el genoma.

Los datos genotípicos de las poblaciones ancestrales utilizados para estimar las proporciones de mestizaje fueron cedidos gentilmente por la Dra. Laura Fejerman (UCSF).

El mestizaje individual fue estimado mediante el programa informático Structure v2.3.4 (Falush et al., 2007). La ancestría genética media fue estimada como el porcentaje de ancestría genética individual dentro de cada grupo. Todos los MIAs estaban en equilibrio de Hardy-Weinberg.

Por otro lado, se analizaron 147 SNPs seleccionados de un panel de genes candidatos y regiones intergénicas asociadas a cáncer diseñado por la Dra. Carolina Bonilla (Oxford University). Luego del genotipado, seis SNPs presentaban un único genotipo en todas las muestras analizadas o presentaron un éxito de genotipado menor al 80%, por lo cual fueron eliminados de los posteriores análisis del estudio. Por otra parte, se eliminaron aquellas muestras en las que no pudo ser genotipado el 10% de los SNPs del panel seleccionado. Se verificó la localización genómica y las frecuencias de poblaciones utilizando dbSNP los genotipos en las distintas (http://www.ncbi.nlm.nih.gov/projects/SNP/) V UCSC Genome Browser (http://genome.ucsc.edu/).

Los ensayos de genotipado de los MIAs y algunos de los SNPs en genes candidatos fueron genotipados por KASPar SNP Genotyping System en KBioscience (Hertfordshire, UK). Los restantes SNPs fueron genotipados por la Dra. Carolina Bonilla en Oxford University.

# 3.2.3 Análisis estadísticos

Una vez obtenido los datos de metilación genómica global, se verificó si tanto en pacientes como en controles presentaban una distribución normal mediante prueba estadística Shapiro-Wilk. La variable metilación fue transformada en logaritmo natural y potencia para lograr ajuste a la distribución normal. En ninguno de los casos los valores de metilación en ambos grupos de muestras presentaron distribución normal, por lo que los siguientes análisis estadísticos se basaron en pruebas no paramétricas.

Para determinar la significancia estadística entre las diferencias encontradas en los
conjuntos de datos de metilación correspondientes a pacientes y controles aplicamos la prueba de Wilcoxon Rank Sum. Dicho prueba permite comparar una variable continua en dos grupos de muestras sin asumir distribución normal. La identificación de independencia entre muestras de individuos afectados e individuos controles para factores de más de 2 estados se realizó aplicando la prueba de Kruskall-Wallis.

Se analizaron los efectos de algunas características propias de los individuos como edad, sexo, ancestría genética, IMC y fumar. Para comparar las características de los casos y controles, utilizamos las pruebas de t de Student para variables continuas y de Fisher para variables categóricas.

Por otro lado, utilizamos regresión logística para los datos de metilación global ajustados por edad y ancestría. La asociación entre metilación del ADN y ancestría fue medida mediante correlación de Kendall Rank. Para visualizar la relación entre ancestría genética y metilación del ADN aplicamos árboles de clasificación y regresión (CART) (Breiman et al., 1984).

La asociación entre los valores de metilación global del ADN de leucocitos y SNPs en genes candidatos de cáncer se estudió mediante prueba Kruskal-Wallis y modelo de regresión lineal. Se aplicó la prueba de Cochran-Armitage, que permite estudiar la asociación entre un locus bialélico y una variable binaria, para analizar la existencia de asociación entre los genotipos de los SNPs y el estatus de la muestra. La comparación de los niveles de metilación entre los distintos genotipos de un SNP determinado fue realizada aplicando prueba de Wilcoxon Rank sum. P valores < 0.05 sin corregir por múltiples ensayos fueron considerados estadísticamente significativos.

Todos los análisis estadísticos fueron realizados utilizando *R environment for statistical computing version 2.15.3* (www.R-project.org) (R Development Core Team, 2013).

## **3.3 Resultados**

# 3.3.1 Hipometilación del ADN genómico de leucocitos en pacientes con cáncer de mama y melanoma

Con el fin de investigar el nivel de metilación global del ADN genómico de leucocitos en pacientes con melanoma cutáneo y cáncer de mama en comparación a individuos controles no afectados, medimos los niveles de metilación genómica en leucocitos mediante cuantificación relativa de 5mdC.

Se midió la concentración relativa de 5mdC en un total de 335 muestras de ADN extraído de leucocitos correspondientes a: 92 pacientes y 113 mujeres controles para el estudio de cáncer de mama, y 65 pacientes y 65 controles para el estudio de melanoma. Luego de eliminar las muestras cuya calidad no fue suficiente para obtener buena resolución en el HPLC o cuyos duplicados diferían en contenido 5mdC más del 3%, se obtuvieron datos de metilación global del ADN en 86/92 (pacientes y controles, respectivamente) en el estudio de cáncer de mama, así como en 42/46 (pacientes y controles, respectivamente) en el estudio de melanoma.

Como se observa en la Tabla 1, al analizar las características epidemiológicas evaluadas en los individuos con datos de metilación global del ADN, detectamos un sesgo en el valor medio de la edad entre pacientes y controles en ambos estudios. En el caso de melanoma, los pacientes son en promedio más jóvenes que los controles. Lo contrario sucede en el estudio de cáncer de mama. Estas diferencias significativas con respecto a la edad deben ser consideradas como posibles confusores en los análisis de metilación del ADN. Con respecto al IMC y el estatus de fumador evaluados en el estudio de cáncer de mama, no existen diferencias significativas entre pacientes y controles. En el estudio de melanoma no se analizaron los datos de IMC ni el estatus de fumador ya que estas dos variables no estaban incluidas en la encuesta realizada a los participantes por no estar descritos como factores importantes de riesgo para la patología.

	Ν	Casos	Ν	Controles	p valor
Edad					
Melanoma	42	57 (24-82)	46	65 (31-82)	0.0110 <sup>c</sup>
Cáncer de mama	86	58 (36-85)	92	53 (36-72)	0.0007 <sup>c</sup>
Sexo (%) <sup>a</sup>	41		45		
Femenino		53.7		64.4	0 3811 <sup>d</sup>
Masculino		46.3		35.6	0.0011
IMC (%) <sup>b</sup>	69		87		
<25		36.2		39.1	0.9186 <sup>d</sup>
25-29,9		34.8		32.2	
>30		29.0		28.7	
Estatus fumador (%) <sup>b</sup>	81		89		
Fumadora		42.0		50.5	0.4599 <sup>d</sup>
No fumadora		58.0		49.5	

**Tabla 1.** Características epidemiológicas evaluadas en los pacientes con cáncer y los individuos controles con datos de metilación global del ADN genómico. Se muestran los valores medios obtenidos en cada categoría.

<sup>a</sup> Estudio caso-control en melanoma

<sup>b</sup> Índice de masa corporal y status de fumador solo evaluado en estudio de cáncer de mama.

<sup>c</sup> Prueba de Student p-valor

<sup>d</sup> Prueba de Fisher p-valor

Con respecto al análisis de metilación global del ADN, tanto en el estudio de cáncer de mama como de melanoma cutáneo encontramos una diferencia significativa en la metilación global del ADN genómico de leucocitos entre individuos con cáncer y sus correspondientes controles no afectados (prueba Wilcoxon Rank Sum, p<0.001; Figura 11 y Tabla 2). El nivel medio de metilación en pacientes con melanoma y cáncer de mama fue significativamente menor ( $2.54 \pm 0.37\%$  y  $2.33 \pm 0.48\%$ , respectivamente) que el nivel medio en los controles ( $2.79 \pm 0.27\%$  y  $2.76 \pm 0.77\%$ , respectivamente). Esta diferencia continúo siendo significativa cuando los datos fueron ajustados por edad y ancestralidad genética mediante regresión logística (melanoma: p=7.10 x  $10^{-4}$ ; cáncer de mama: p= $5.47 \times 10^{-4}$ ). Al comparar la distribuciones de los valores de metilación en los pacientes y en los controles (Figura 12). En el estudio de cáncer de mama, en el grupo de pacientes prácticamente no se detectan niveles de metilación superiores a 3%, mientras que los controles presentan una distribución hacia niveles más elevados. De la

misma manera, en el estudio de melanoma el grupo control presenta mayor frecuencia de niveles elevados de metilación global que el grupo de pacientes.



**Figura 11.** Comparación de los niveles de metilación global del ADN genómico en leucocitos de pacientes con cáncer e individuos controles. El estudio caso-control de cáncer de mama se muestra en rojo, el estudio caso-control de melanoma cutáneo se muestra en azul. Los "*boxes*" representan el rango intercuartil y la línea horizontal indica la media. Las diferencias estadísticamente significativas entre pacientes con cáncer y controles fueron determinadas utilizando la prueba Wilcoxon Rank Sum (\* = p<0.001).

	Ν	Casos	Ν	Controles	p valor
Metilación global del ADN					
Melanoma <sup>ª</sup> Cáncer de mama <sup>ª</sup>	42 86	2.542 % 2.330 %	46 92	2.794 % 2.769 %	9.96 x 10 <sup>-4</sup> 5.96 x 10 <sup>-5</sup>
Metilación global del ADN-Sexo					
Melanoma Mujeres Hombres	22 20	2.480 % 2.628 %	29 17	2.793 % 2.805 %	0.269 <sup>b</sup> 0.198 <sup>c</sup>

**Tabla 2.** Media de la metilación global del ADN genómico de leucocitos en pacientes con cáncer y controles.

<sup>a</sup> Regresión logística para melanoma p=7.10 x 10<sup>-4</sup> y cáncer de mama p=5.47 x 10<sup>-4</sup>, ajustado por ancestralidad y edad.

<sup>b</sup> Comparación entre mujeres y hombres en los casos.

<sup>c</sup> Comparación entre mujeres y hombres en los controles.

Esta forma de visualizar los datos, corrobora nuevamente la presencia de hipometilación global del ADN de leucocitos en los grupos de pacientes con cáncer en comparación a sus respectivos grupos controles, donde la condición tumoral tiende a minimizar incluso las diferencias intrapoblacionales.



**Figura 12.** Distribución de los niveles de metilación global del ADN en cada individuo del estudio. Frecuencia de los niveles de metilación a lo largo de **(A)** 86 pacientes con cáncer de mama y 92 controles, y **(B)** 42 pacientes con melanoma cutáneo y 46 controles.

Para evaluar el efecto de potenciales confusores sobre la metilación global del ADN de leucocitos, examinamos la asociación del estatus de la enfermedad (paciente/control) con la edad, sexo, estatus de fumador e IMC. Dado que encontramos evidencias de diferencias significativas solo en la edad entre casos y controles (Tabla 1), analizamos la correlación entre los niveles de metilación global del ADN genómico de leucocitos y la edad tanto en el estudio de melanoma como en el de cáncer de mama, y no detectamos una asociación significativa (Prueba de correlación Kendall, p > 0.05, Tabla 4). Tampoco evidenciamos asociación entre el porcentaje de metilación del ADN genómico e IMC o estatus de fumador en el estudio de cáncer de mama (Kruskal Wallis, p >0.1). En el caso de melanoma, no detectamos diferencias significativas en los niveles de metilación global del ADN en función del género en casos y controles (Wilcoxon Rank Sum, p > 0.05) (Tabla 2).

La hipometilación global del ADN genómico detectada en leucocitos de los pacientes con cáncer con respecto a los individuos libres de patología demuestra la existencia de marcas diferenciales y sugiere que los patrones de metilación del ADN analizados en sangre periférica tienen el potencial de ser biomarcadores informativos de riesgo a melanoma y cáncer de mama.

# 3.3.2 Ancestría genética en la metilación global del ADN genómico de leucocitos de pacientes con cáncer

Debido a que Uruguay tiene una población mestizada y tanto el cáncer de mama como melanoma presentan altas incidencias en europeos (Globocan 2012, www.globocan.iarc.fr), decidimos investigar la influencia de la ancestría genética en los niveles de metilación global del ADN de leucocitos en pacientes de ambos tipos de cánceres.

Previamente en nuestro laboratorio contábamos con los datos de mestizaje individual de todos los participantes de ambos estudios caso-control. El grado de mestizaje de todos los individuos fue determinado analizando un grupo de 59 SNPs que pueden identificar ancestralidad nativo americana, africana y europea. En el análisis de mestizaje, primero estudiamos la distribución de la ancestría genética a lo largo de los individuos de cada categoría a comparar (pacientes versus controles), no encontrándose diferencias significativas en la contribución ancestral individual tanto en el estudio de melanoma como de cáncer de mama (Tabla 3).

Sin embargo, si consideramos todos los individuos como un único conjunto, detectamos mayores porcentajes de contribución ancestral europea en el estudio de melanoma que en el estudio de cáncer de mama. Este sesgo entre ambos estudios podría ser atribuido a diferencias en los lugares de muestreo, ya que en el caso de melanoma todas las muestras participantes del proyecto fueron recolectadas únicamente en Montevideo, mientras que en el proyecto de cáncer de mama el muestreo cubre varias regiones del país (ver sección 2.3.2).

	Ν	Casos	Ν	Controles	p valor
Melanoma	49		73		
Europeo		95.04 %		93.32 %	0.184
Nativo Americano		3.33 %		4.94 %	0.211
Africano		1.16 %		1.74 %	0.120
Cáncer de mama	179		209		
Europeo		76.89 %		76.49 %	0.927
Nativo Americano		12.86 %		13.85 %	0.229
Africano		10.25 %		9.66 %	0.420
· · · · ·				· · ·	

**Tabla 3.** Comparación del componente genético ancestral individual entre pacientes con cáncer

 y controles no afectados. Se indican los valores medios de cada categoría.

Con el objetivo de determinar si la ancestría individual tiene influencia en los niveles de metilación genómica global, llevamos adelante un análisis de correlación entre el porcentaje de contribución africana, europea y nativa americana, y el de metilación global del ADN en leucocitos. Como se muestra en la Tabla 4, el análisis de correlación de Kendall revela una asociación inversa significativa entre el componente ancestral africano y el porcentaje de metilación global del ADN de leucocitos en pacientes con cáncer de mama ( $\tau = -0.199$ , p<0.01). Por otro lado, no observamos asociación significativa entre la metilación y la ancestría genética en pacientes con melanoma, probablemente debido al tamaño reducido de la muestra con ambos tipos de datos. Sin embargo, cuando consideramos todos los pacientes de cáncer de mama y melanoma juntos, observamos que la correlación negativa con el componente ancestral africano se vuelve más fuerte ( $\tau = -0.187$ ; p < 0.005; Tabla 4 y Figura 13a). De igual modo, también encontramos una correlación positiva significativa con el componente ancestral europeo ( $\tau = 0.169$ ; p < 0.01; Tabla 4 y Figura 13b). Contrariamente a estos resultados, no pudimos detectar ninguna correlación entre el nivel de metilación del ADN y la ancestría genética en los controles.

	N <sup>a</sup>	Edad	N <sup>a</sup>	Europeo	Nati∨o americano	Africano
Melanoma	86	-0,023				
Casos Controles	41 45	-0.140 -0.034	28 39	0.184 -0.125	-0.204 0.153	-0.132 0.000
Cáncer de mama	169	-0.028				
Casos Controles	80 89	-0.008 0.064	78 89	0.124 -0.024	-0.062 0.080	-0.199 <sup>b</sup> -0.074
Muestreo total						
Casos Controles	121 134	-0.049 0.072	106 128	0.169 <sup>b</sup> 0.017	-0.127 0.045	-0.187 <sup>c</sup> -0.098

**Tabla 4.** Coeficientes de correlación de Kendall y sus significancias estadísticas entre metilación global del ADN de leucocitos, edad y componentes individuales ancestrales.

<sup>a</sup>Los individuos sin datos de edad y/o ancestría fueron removidos de este análisis. <sup>b</sup>p <0.01 <sup>c</sup>p <0.005



**Figura 13.** Correlación entre la ancestría genética individual y la metilación global del ADN en leucocitos de pacientes con cáncer. **(a)** El componente ancestral africano se correlaciona negativamente con la metilación del ADN (r = -0.187, p<0.005). **(b)** El componente ancestral europeo se correlaciona positivamente con la metilación del ADN (r = -0.187, p<0.005). **(b)** El componente ancestral europeo se correlaciona positivamente con la metilación del ADN (r = 0.169, p<0.01).

La influencia de la ancestría genética en los niveles de metilación del ADN genómico de leucocitos de pacientes con cáncer puede ser claramente visualizada mediante un árbol de regresión y clasificación (CART) del estatus de la enfermedad en los niveles de metilación del ADN ajustado por edad. Esta representación muestra los valores o porcentajes predictivos de metilación del ADN dependiendo de la frecuencia de los componentes africano y europeo en cada paciente con cáncer (Figura 14). Como se visualiza en la figura, cuanto mayor es el componente genético africano, menor es el porcentaje de metilación del ADN de leucocitos detectado en el paciente con cáncer.



**Figura 14.** Análisis de árboles de clasificación y regresión (CART) de metilación global del ADN de leucocitos y componente genético ancestral en pacientes con cáncer. La ancestría se muestra como frecuencia, y los porcentajes predictivos de metilación global del ADN se muestran en los extremos de las ramas.

# 3.3.3 Asociación de genotipos en genes candidatos y metilación global del ADN genómico en pacientes con cáncer de mama

Habiendo demostrado que el contexto genético del paciente con cáncer, dado por su componente genético ancestral, influye en los niveles de metilación global en sangre, decidimos continuar integrando la información epigenética y genética de los individuos.

En nuestro laboratorio se realizaron estudios de asociación de genes candidatos de cáncer y el riesgo a cáncer de mama en nuestra población. Se contaba con los datos genotípicos de 141 SNPs en todos los pacientes con cáncer de mama y sus respectivos controles. Utilizando estos datos nos planteamos estudiar polimorfismos genéticos en genes candidatos de cáncer y su posible asociación con el nivel de metilación genómica global. Varios de los polimorfismos analizados fueron reportados previamente no solo asociados a cáncer sino también involucrados directa o indirectamente en alteraciones de los procesos de metilación del ADN (por ejemplo: *MTHFR, DNMT3A, BRCA1, DNMT3B*) (Tabla 5).

En el estudio de cáncer de mama esporádico se estudió la asociación entre los valores de metilación global del ADN de leucocitos y los genotipos de 141 SNPs ubicados en 95 genes candidatos de cáncer y 3 regiones intergénicas, distribuidos en los 22 cromosomas autosómicos. El análisis inicial se realizó sin tener en cuenta el estatus de las muestras (paciente/control), detectándose asociación significativa entre el porcentaje de metilación global y los genotipos de los siguientes 8 SNPs (prueba Kruskal Wallis, p < 0.05) (Tabla 5): rs3819954 (gen *TP73*), rs1801270 (gen *CDKN1A*), rs7781370 (gen *FLJ42280*), rs3817198 (gen *LSP1*), rs1007738 (gen *CKAP5*), rs545500 (gen *MUS81*), rs7191944 (gen *TUBB3*) y rs1006899 (región intergénica en cromosoma 21).

**Tabla 5.** Asociación entre SNPs en genes candidatos de cáncer y niveles de metilación genómica global en leucocitos de individuos del estudio caso-control de cáncer de mama. En rojo se indican las asociaciones significativas. Cr: cromosoma.

SNPs	Cr	Gen	p valor	SNPs	Cr	Gen	p valor
rs3819954	1	TP73	0.02972	rs3130340	6	MHC/C6orf10	0.07447
rs1801133	1	MTHFR	0.9819	rs3176336	6	CDKN1A	0.3908
rs6426749	1	ZBTB40	0.6382	rs1801270	6	CDKN1A	0.02668
rs1430742	1	GPR177	0.3873	rs2479717	6	CCND3	0.402
rs2297138	1	CYR61	0.8085	rs4870044	6	C6orf97	0.917
rs1571549	1	CYR61	0.5446	rs2941740	6	ESR1	0.3783
rs17131547	1	TGFBR3	0.4773	rs1999805	6	ESR1	0.7389
rs1047840	1	EXO1	0.07019	rs1514347	6	ESR1	0.9258
rs4665777	2	DNMT3A	0.1652	rs1062577	6	ESR1	0.6489
rs87938	3	CTNNB1	0.7301	rs2273206	6	ESR1	0.5676
rs4858795	3	ATRIP/SCOTIN	0.5471	rs2982712	6	ESR1	0.6043
rs13059232	3	NR112	0.565	rs6912184	6	ESR1	0.0631
rs361072	3	PIK3CB	0.1209	rs827423	6	ESR1	0.7244
rs28897764	3	ATR	0.7787	rs1524058	7	STARD3NL	0.4053
rs2056116	4		0.9254	rs2708861	7	HUS1	0.5402
rs3736599	4	SULT1E1	0.4664	rs7781370	7	FLJ42280	0.04131
rs1471403	4	MEPE	0.2034	rs884716	7	DSS1/SHFM1	0.3882
rs13123099	4	ADH1A	0.1727	rs2214681	7	CNTNA P2	0.6686
rs2255137	4	hCDC4/FBXW7	0.3408	rs1805794	8	NBS1/NBN	0.3828
rs981782	5	HCN	0.1842	rs1805841	8	NBS1/NBN	0.07861
rs16886165	5	MAP3K1	0.4129	rs2073618	8	TNFRSF11B	0.8172
rs889312	5	MAP3K1	0.7696	rs6993813	8	OPG	0.644
rs1366594	5	MEF2C	0.3323	rs3731239	9	CDKN2A	0.739

SNPs	Cr	Gen	p valor	SNPs	Cr	Gen	p valor
rs2855192	9	ABL1	0.3034	rs175080	14	MLH3	0.533
rs12762549	10	CYP2D6	0.4719	rs1885602	14	GTF2A1	0.252
rs1219648	10	FGFR2	0.7203	rs4899786	14	GTF2A1	0.4287
rs2981582	10	FGFR2	0.5415	rs1952585	14	ESR2	0.2516
rs3817198	11	LSP1	0.04193	rs3020449	14	ESR2	0.6361
rs7117858	11	SOX6	0.3396	rs7154455	14	ESR2	0.4052
rs7932354	11	ARHGAP1	0.539	rs2412546	15	RAD51	0.6535
rs1007738	11	CKAP5	0.03424	rs1876206	15	FBN1	0.4801
rs1503185	11	PTPRJ	0.3213	rs951715	15	IGF1R	0.1079
rs545500	11	MUS81	0.01336	rs3803662	16	TOX3/LOC643714	0.09807
rs599083	11	LRP5	0.2355	rs16260	16	CDH1	0.1952
rs9344	11	CCND1	0.7052	rs9929218	16	CDH1	0.1789
rs556477	11	MRE11A	0.1720	rs11649550	16	FXBO31	0.6035
rs1799750	11	MMP1	0.2651	rs4843599	16	FXBO31	0.3180
rs3218707	11	ATM	0.1757	rs9937099	16	FXBO31	0.4494
rs4987945	11	ATM	0.751	rs3212363	16	MC1R	0.4554
rs1800056	11	ATM	0.3515	rs7191944	16	TUBB3	0.04560
rs3217901	12	CCND2	0.8562	rs1799941	17	SHBG	0.4734
rs1042725	12	HMGA2	0.6947	rs16942	17	BRCA1	0.7334
rs2172881	12	GRIP1	0.843	rs1799950	17	BRCA1	0.731
rs2279744	12	MDM2	0.3911	rs8176092	17	BRCA1	0.6725
rs8176882	12	PAWR	0.7766	rs8176193	17	BRCA1	0.5706
rs12827748	12	PAWR/LOC100133105	0.3745	rs1513670	17	SOST	0.7034
rs7136446	12	IGF1	0.6528	rs9303521	17	CRHR1	0.3608
rs35009941	12	TXNRD1	0.953	rs1042522	17	TP53	0.1758
rs7136874	12	BRAP	0.3651	rs2078486	17	TP53	0.8292
rs649406	12	BRAP2/ACAD10	0.7366	rs8064946	17	TP53	0.8769
rs747443	12	NCOR2	0.4876	rs1978503	18		0.1976
rs864503	12	NCOR2	0.4632	rs884205	18	TNFRSF11A	0.05167
rs2238136	12	VDR	0.606	rs3018362	18	RANK	0.3228
rs2239179	12	VDR	0.7627	rs2279115	18	BCL2	0.8223
rs2239182	12	VDR	0.6415	rs713041	19	GPX4	0.5366
rs4237855	12	VDR	0.4569	rs1056538	19	ICAM5	0.793
rs7299460	12	VDR	0.5096	rs11670365	19	CA RM1	0.3820
rs7975232	12	VDR	0.3149	rs997669	19	CCNE1	0.4625
rs731236	12	VDR	0.1006	rs406193	20	DNMT3B	0.4472
rs1544410	12	VDR	0.07125	rs3918242	20	MMP9	0.5219
rs206143	13	BRCA2/LOC646799	0.6912	rs8183919	20	PTGIS	0.7249
rs9533090	13	AKAP11	0.7876	rs6064389	20	CSTF1	0.1101
rs9594759	13	RANKL	0.3945	rs1006899	21		0.01253
rs2854344	13	RB1	0.6579	rs926184	21	NRIP1	0.481
rs766173	13	BRCA2	0.7399	rs458685	21	GRIK1	0.9895
rs4987117	13	BRCA2	0.5942	rs17881274	21	SOD1	0.609
rs1799954	13	BRCA2	0.1056	rs4680	22	COMT	0.8577
rs144848	13	BRCA2	0.2373	rs933226	22	YWHAH	0.4430
rs206081	13	BRCA2	0.7668	rs20551	22	EP300	0.8881
rs4942505	13	BRCA2	0.4954	rs11571787	22	CYP2D26	0.3815
rs9534174	13	BRCA2	0.2854				

Tabla 5. Continuación

Se verificó que estos SNPs no tenían asociación entre sus genotipos y la pertenencia de las muestras al grupo de pacientes o controles, descartando la posibilidad de que la asociación con los niveles de metilación se deba a diferencias en la

distribución de los genotipos entre pacientes y controles.

Con el fin de evaluar el efecto potencial de los polimorfismos en la metilación global del ADN, comparamos los niveles de metilación global del ADN entre los diferentes genotipos de los 8 SNPs de manera independiente mediante análisis de regresión lineal ajustado por estatus de la muestras (paciente/control) y edad. Se observaron diferencias significativas en los niveles de metilación global entre los genotipos de los SNPs en *TP73* (p = 0.0007), *FLJ42280* (p = 0.0001), *MUS81* (p = 2.43 x 10<sup>-05</sup>) y el rs1006899 en región intergénica del cromosoma 21 (p = 0.0008).

El modelo de asociación entre el SNPs rs545500 (*MUS81*) y la metilación global fue el que se ajustó con mayor significancia estadística ( $p = 2,43 \times 10^{-05}$ ). Evaluamos el cambio aminoacídico generado por este SNPs con los programas SIFT Human SNPs (http://sift.jcvi.org/) y PolyPhen-2 (//genetics.bwh.harvard.edu/pph2/), y no tendría consecuencias sobre la estructura de la proteína. Sin embargo, como veremos más adelante, el gen *MUS81* está involucrado en los mecanismos de reparación del ADN e indirectamente podría estar también involucrado en mecanismos epigenéticos. Por ello, continuamos evaluando este SNPs en mayor profundidad.

De acuerdo al análisis de regresión, el efecto de los 3 genotipos del SNPs en *MUS81* es diferente en pacientes y en controles (Tabla Suplementaria 3). Los individuos controles con el genotipo CC del SNPs en *MUS81* presentan niveles de metilación genómica global más alto que los individuos con los genotipos alternativos CG/GG del mismo locus (Wilcoxon test, p = 0,0018). Por el contrario, en los pacientes no se observan diferencias significativas en los niveles de metilación entre individuos con genotipo CC e individuos con genotipos CG/GG (Wilcoxon test, p > 0,05), aunque en el modelo predictivo del análisis de regresión se marca una tendencia a valores mayores de metilación global en pacientes homocigota para el alelo C (Tabla Suplementaria 3).

Cabe destacar que no encontramos asociación con SNPs ubicados en los genes codificadores de la ADN metiltransferasas o de genes involucrados en el metabolismo de los grupos dadores de metilos, los cuales están directamente involucrados en los mecanismos epigenéticos.

## 3.4 Discusión

Hemos demostrado la existencia de hipometilación global del ADN en leucocitos y una correlación negativa entre la ancestría genética individual africana y la metilación global del ADN en pacientes con cáncer esporádico. Hasta el momento, no existen estudios previos que hayan investigado la influencia de la ancestría genética individual en la metilación global del ADN de leucocitos en pacientes con cáncer.

# 3.4.1 Hipometilación global del ADN de leucocitos como biomarcador de melanoma y cáncer de mama esporádico

Se ha propuesto que la hipometilación genómica global tiene un importante impacto en la biología tumoral mediante la generación de inestabilidad cromosómica, reactivación de elementos transponibles, y pérdida de impronta genómica (Esteller, 2008). Existe evidencia creciente de que los leucocitos son un modelo celular útil para evaluar cambios epigenéticos. De manera que las epimutaciones y la hipometilación global del ADN, asociada a un aumento del riesgo a cáncer, podrían ser detectadas en sangre periférica en lugar del tejido afectado (Woo and Kim, 2012). Esto es de gran importancia ya que las muestras de sangre son fáciles de obtener mediante técnicas no invasivas, siendo muy útil para estudios epidemiológicos de gran escala (Li et al., 2012).

En este trabajo por primera vez se demuestra una asociación significativa de hipometilación global del ADN de leucocitos con melanoma cutáneo maligno esporádico. Nuestros resultados son contrarios a un trabajo previo en melanoma donde no se encontró asociación entre la metilación global o CpG específica de LINE-1 en sangre periférica y melanoma cutáneo maligno familiar (Hyland et al., 2013). Sin embargo, existen dos grandes diferencias entre ambos estudios. Primero, los niveles de metilación en un elemento repetitivo (LINE-1) no necesariamente representa el contenido de 5-metilcitosina a lo largo de todo el genoma, ya que existen muchos otros contextos genómicos que se encuentran metilados (secuencias Alu, SINE, genes improntados). Segundo, el trabajo de Hyland y colaboradores (2013) examina la

asociación entre metilación del ADN y melanoma cutáneo en familias tendientes a este tipo de cáncer, sugiriendo que en este caso el mecanismo genético es más importante que el efecto epigenético. De manera interesante, varios trabajos han demostrado una hipometilación global en líneas celulares de melanoma (Tellez et al., 2009), cultivos *in vitro* de tumores de melanoma cutáneo extirpados quirúrgicamente (Sigalotti et al., 2011) y tumores primarios de melanoma (Ecsedi et al., 2013). Pero todos ellos evaluando la metilación global en CpG de secuencias repetidas LINE-1 y sobre el tejido afectado o líneas celulares, demostrándose la importancia de hipometilación en LINE-1 como marcador pronóstico y la asociación de su estatus de metilación con la capacidad de metástasis del tumor.

Por otro lado, en el estudio caso control de cáncer de mama corroboramos los hallazgos de otros autores en relación a la asociación de la hipometilación del ADN en sangre periférica con el riesgo a cáncer de mama (Choi et al., 2009; Delgado-Cruzata et al., 2012; Deroo et al., 2014). Por ejemplo, un estudio basado en casos-controles hospitalarios de 180 pacientes con cáncer de mama y 180 controles no afectados muestra una disminución de 5mdC, medida por espectrometría de masa, en el ADN de leucocitos de los casos con cáncer (Choi et al., 2009). El nivel de metilación global del ADN encontrado en nuestro estudio es levemente menor que lo reportado en el estudio de Choi y colaboradores (2009). Esta diferencia podría deberse al uso de diferentes protocolos y métodos con distinta sensibilidad para medir la metilación genómica global. Por ello, un desafío para que la hipometilación global sea un marcador clínico útil será definir y normalizar las metodologías para su determinación en gran número de pacientes. Además, hay que tener en cuenta que el nivel global de metilación del ADN reportado previamente en muestras de sangre varía sustancialmente (2,3 a 6 %) dependiendo de la enfermedad estudiada o si es referida a una población sana, del rango de edad de las muestras y de los factores ambientales (Fraga et al. 2005; Lim et al. 2008; Moore et al. 2008; Fuke et al. 2004; Kok et al. 2007; Ma et al. 2009; Guerrero-Preston et al. 2007). Como discutiremos más adelante, la novedad de nuestro estudio con respecto a los resultados obtenidos previamente por Choi y colaboradores (2009) radica en que además analizamos la asociación de los niveles de metilación global del ADN detectada en leucocitos con el contexto genético de los individuos: ancestralidad genética individual y polimorfismos genéticos.

Gran cantidad de evidencia apoya el hecho de que la metilación del ADN a nivel global, sobre todo en la sangre, cambia durante el envejecimiento, algunas características fenotípicas y estilos de vida, como la edad, sexo, índice de masa corporal, la dieta y el tabaquismo (Terry et al., 2011; Zhu et al., 2012b). Sin embargo, los niveles de metilación genómica global no se asociaron con ninguna de las variables fenotípicas y de estilo de vida evaluados. En este trabajo la mayoría de los participantes son de edad avanzada (media= 55.5 años en cáncer de mama; media= 61 años en melanoma). Esto puede explicar la pérdida de asociación entre la metilación del ADN y la edad en nuestra población de estudio, en contraste a otros estudios que incluyen individuos con un rango de edad más amplio (Fraga et al., 2005). Tampoco se analizó la dieta ni todos los factores genéticos involucrados en el metabolismo del carbono-1 reportados como posibles modificadores de los patrones de metilación del ADN. El metabolismo de carbono-1 mediado por folato es una red de vías metabólicas interconectadas necesarias para la síntesis de nucleótidos de purina, la timidilato y la remetilación de la homocisteína a metionina. Las alteraciones en esta vía, como resultado de deficiencias nutricionales y variantes genéticas comunes, influyen en la síntesis de ADN y la estabilidad y metilación de la cromatina (Stover, 2011). Por ello analizamos la posible asociación entre el SNPs rs1801133, correspondiente a la variante C677T en el gen MTHFR (5,10-metilen tetrahidrofolato reductasa), y los niveles de metilación global en los individuos del estudio caso-control de cáncer de mama. Esta variante es la mutación más frecuente en la población uruguaya en los genes involucrados en el metabolismo del carbono-1, con una frecuencia de 8% de homocigotas (Otero et al., 2004). Sin embargo, no encontramos ninguna asociación de dicha variante con la metilación global en sangre. Si bien no se evaluó la dieta ni el resto de las variantes genéticas comunes en el metabolismo del carbono-1 como modificadores en nuestro estudio, las asociaciones reportadas en estudios humanos con respecto a estas variables y la metilación global del ADN son aún inconsistentes (Gomes et al., 2012; Moore et al., 2008; Narayanan et al., 2004; Terry et al., 2011). Tampoco podemos descartar el efecto que ejerce la condición tumoral. En situaciones normales (individuos sanos) factores como la edad o la dieta pueden contribuir a cambios pequeños en los perfiles de metilación del ADN. El cáncer es una situación patológica que genera un gran número de alteraciones en el individuo, que más allá de la acumulación de mutaciones genéticas y epigenéticas en el tejido tumoral, también se producen cambios en el metabolismo general del individuo, procesos inflamatorios,

entre otros. Por ello, es probable que el efecto de la patología enmascare los otros efectos ambientales.

Aun no hay evidencias claras si existe correlación entre la metilación del ADN en sangre periférica y el tejido tumoral, por lo que sigue siendo controversial en cuanto a si la metilación del ADN de leucocitos de sangre periférica refleja la metilación del tejido diana (Brennan and Flanagan, 2012). Se ha sugerido que los procesos inmunológicos relacionados con la inflamación en el desarrollo del cáncer conducen a cambios en las subpoblaciones de leucocitos, que pudiera alterar las firmas epigenéticas en el ADN de sangre periférica (Teschendorff et al., 2009). Otra posible explicación es que el cambio epigenético debido a la metilación se asocia con las variantes genéticas de cánceres específicos (Pedersen et al. 2011; Teschendorff et al. 2009).

Estudios recientes sobre la heredabilidad de los patrones de metilación del ADN apoyan el postulado de que los factores genéticos son determinantes importantes. En un estudio de adultos sanos examinados en dos momentos con más de una década de diferencia, algunos individuos perdieron pero otros ganaron metilación del ADN y los patrones de cambio se agruparon en familias, lo que sugiere que la variación genética puede influir en los patrones de metilación (Bjornsson et al., 2008).

En resumen, detectamos metilación global del ADN genómico en los leucocitos de sangre periférica significativamente diferente entre los controles sanos y pacientes con cáncer de mama y pacientes con melanoma. Estos resultados apoyan nuestra hipótesis de que algunos cambios epigenéticos sistémicos pueden ser detectados en el ADN de sangre periférica en pacientes con cáncer esporádico.

Es necesario incrementar el conocimiento del origen y la naturaleza de la metilación del ADN en leucocitos de sangre periférica para determinar si la metilación del ADN en estas células puede servir como biomarcador informativo de susceptibilidad al desarrollo tumoral o incluso como marcador temprano de inicio del cáncer.

# 3.4.2 El contexto genético sobre la variabilidad de la metilación global del ADN

#### Ancestría genética individual de pacientes con cáncer

Las poblaciones de los distintos continentes varían considerablemente en su predisposición a las enfermedades, probablemente como resultado de procesos adaptativos a factores selectivos ambientales locales y posterior cambio de hábitos de vida. Ejemplo de ello son el cáncer de mama y el melanoma cutáneo maligno, los cuales presentan diferencias étnicas en cuanto a su incidencia y mortalidad (Globocan 2012, www.globocan.iarc.fr).

Estudiamos por primera vez el efecto de la ancestría genética individual en la metilación global del ADN en leucocitos en pacientes con cáncer de mama o melanoma cutáneo esporádico de una población mestizada. La población uruguaya, dado su calidad de población mestizada tri-híbrida, es idónea para el estudio de los efectos de los distintos componentes genéticos ancestrales en una enfermedad compleja como el cáncer.

Recientemente se plantea que los patrones de metilación del ADN varían entre las distintas poblaciones, incluso entre los distintos grupos étnicos de una misma región. Se ha detectado la existencia de diferencias étnicas leves al nacer en sitios CpGs específicos, y el patrón predominante es el de menores niveles de metilación en CpG en niños recién nacidos afroamericanos (Adkins et al., 2011). De manera similar, se describió menor nivel de metilación del ADN en mujeres adultas afroamericanas en relación a mujeres europeo americanas (Terry et al., 2008). En el presente trabajo encontramos una correlación negativa entre el componente ancestral africano y la metilación genómica global en leucocitos de pacientes con cáncer. Asimismo, no detectamos diferencia entre los pacientes y los controles con respecto a la ascendencia genética individual. Por lo tanto, la hipometilación del ADN observada en los pacientes no se podría atribuir a las diferencias en la ascendencia genética entre los afectados y controles.

Somos conscientes que dado el tamaño muestral relativamente pequeño de nuestros grupos caso-control, estos resultados deben ser confirmados con estudios mayores y en otras poblaciones mestizadas. Debemos de notar que en ambos estudios los controles

pueden no ser representativos de la población general ya que son controles muestreados en hospitales. No tuvimos acceso a la información del grado tumoral o etapa de la enfermedad en todos los pacientes, por lo que estudios adicionales se requerirán para clarificar si la influencia de la ancestría genética en los niveles de metilación del ADN en los pacientes con cáncer podría tener consecuencias en la agresividad del tumor o en el pronóstico en pacientes con diferente componente ancestral.

Los niveles de metilación en sitios CpG individuales pueden estar fuertemente asociados con variación en secuencias génicas locales y distantes (Gibbs et al. 2010; Zhang et al. 2010). Por otro lado, las frecuencias de los alelos de SNP pueden diferir sustancialmente entre poblaciones con diferentes ascendencias geográficas (Altshuler et al., 2010). Por lo tanto, es posible que las diferencias étnicas observadas en la metilación del ADN puedan deberse a diferencias a nivel de secuencias de regiones cromosómicas que influyen en los niveles de metilación global y CpG específica, y que identificamos por la diferencia de alelos de SNP o haplotipos entre los individuos. Aún sin existir diferencias en ancestría entre los controles y pacientes analizados en nuestro estudio, los últimos tendrían secuencias africanas exclusivas asociadas a hipometilación genómica global en leucocitos. Así, la ancestría genética individual podría contribuir a la variación interindividual de la metilación del ADN entre pacientes con cáncer.

Lo que es más, la frecuencia de las variantes en los genes cuyos productos están implicados en la metilación de ADN difiere entre los grupos étnicos, tales como las variantes en los genes que codifican para ADN metiltransferasa 2 y 3B (Franchina & Kay 2001; Zhao et al. 2009; Zhu et al. 2012). Por ejemplo, la variante 579G>T en el promotor de DNMT3B presenta una distribución diferente entre las poblaciones chinas, coreanas y americanas (Zhao et al., 2009). Algunas de las variantes en las DNMTs podrían alterar la expresión y función de las enzimas implicadas en los mecanismos epigenéticos y ayudarían a explicar las diferencias en el grado de metilación observado entre las poblaciones.

Fraser y colaboradores (2012) sugirieron que la metilación del ADN es altamente divergente entre las poblaciones, y que esta divergencia podría deberse en parte a la combinación de diferencias en frecuencias alélicas y epístasis complejas o interacciones genoma-ambiente. De esta manera, una variante que está presente en dos poblaciones, podría afectar a la metilación del ADN únicamente en una de ellas (Fraser et al., 2012). Por otra parte, Heyn y colaboradores (2013) identificaron diferencias en la metilación del ADN que son capaces de distinguir tres grandes grupos étnicos humanos (eurodescendientes, afrodescendientes y asiáticos) y QTL (quantitative trait loci)-CpG específicos asociados con la variación natural humana, contribuyendo a determinar características fenotípicas diferentes entre las poblaciones humanas. Por lo tanto, la estructura tri-híbrida de la población mestiza uruguaya podría estar parcialmente explicando el patrón de metilación. La ancestría genética afrodescendiente jugaría un rol particular en la susceptibilidad o etiología de estos cánceres, promoviendo una hipometilación genómica sistémica, no necesariamente de la misma manera que en otras poblaciones del mundo.

En conclusión, nuestros hallazgos claramente indican la influencia de la ancestría genética individual en el nivel de metilación global del ADN de leucocitos en pacientes con cáncer de Uruguay, la cual podría resultar en patrones alterados de expresión génica y diferencias en el desarrollo clínico y pronóstico de la enfermedad. Estos resultados resaltan la importancia de tomar en cuenta la ancestría genética cuando se analizan datos epigenéticos a nivel poblacional. Aún más importante, la posibilidad de que la ancestría genética pudiera estar asociada a regiones cromosómicas propensas a la metilación del ADN (o propensas a estar desmetiladas) sugiere que la técnica de mapeo génico por mestizaje (Winkler et al., 2010) podría extenderse a relacionar un segmento cromosómico ancestral, el estado de metilación y la susceptibilidad a la enfermedad. Así, la prevalencia de alteraciones epigenéticas puede proporcionar una base para la comprensión de las diferencias en el inicio temprano del cáncer, su agresividad, y malos pronósticos observados en varias poblaciones étnicas. La identificación de estas diferencias en los procesos epigenéticos en diversas poblaciones podría contribuir al desarrollo de biomarcadores de riesgo de la enfermedad y a un mejor diseño de las intervenciones terapéuticas.

#### Variabilidad genética en genes candidatos de cáncer

La investigación de polimorfismos genéticos que pudieran estar asociados a la metilación global del ADN puede ayudar a dilucidar la diferencia en los niveles de metilación global observados de acuerdo a la ancestría de los individuos. Ejemplos claros de ello son los polimorfismos en los genes codificadores de metiltransferasas del ADN y *MTHFR*, que como se discutió anteriormente, tienen una distribución de

frecuencias alélicas distintas según los grupos étnicos o poblaciones y están directamente asociados a los mecanismos de metilación del ADN. Sin embargo, no detectamos asociación de los polimorfismos analizados en estos genes con los niveles de metilación global del ADN detectados en pacientes con cáncer de mama ni en su grupo control. Por lo que las frecuencias diferenciales de estas variantes génicas entre los individuos no pueden explicar la variación interindividual en la metilación global del ADN observada en este estudio.

Por otro lado, al analizar otros SNPs en regiones candidatas previamente asociadas a cáncer, detectamos una asociación positiva altamente significativa de la variante CC del SNPs rs545500 en el gen *MUS81* con el nivel de metilación global del ADN en los individuos controles de nuestro estudio de cáncer de mama. También detectamos asociación de los genotipos con el nivel de metilación del ADN, sin considerar el estatus de la muestra, en otros 3 SNPs (ubicados en *TP73, FLJ42280* y región intergénica en cromosoma 21). Sin embargo, el efecto de los distintos alelos de *TP73* si bien es significativo, las diferencias sobre la metilación del ADN son muy leves. Los SNPs en el gen putativo *FLJ42280* y en la región intergénica del cromosoma 21 fueron previamente asociados a baja densidad ósea mineral, un factor de riesgo asociado a cáncer de mama (Styrkarsdottir et al., 2009). Pero dado que estos SNPs no se encuentran en genes con función conocida, establecer una relación causal de dicha asociación es dificultosa.

Sin embargo, el SNPs rs545500 es un SNPs no sinónimo localizado en la región codificante del gen *MUS81*. Dicho gen codifica una nucleasa específica de la estructura del ADN que juega un rol importante en la reparación del ADN por recombinación homóloga (Osman and Whitby, 2007). Este polimorfismo resulta en un cambio de aminoácido de arginina (residuo hidrofóbico cargado positivamente) a prolina (residuo hidrofóbico sin carga) en la posición 180 de la proteína, que no tendría consecuencias sobre la estructura de la proteína. Sin embargo, no podemos descartar que se afecte la interacción proteína-proteína fundamental para la correcta función de MUS81 (Nakken et al., 2007).

El rol del gen *MUS81* en el cáncer de mama aún no se ha investigado. Sin embargo, se ha detectado una baja expresión de este gen en tumor gástrico y carcinoma colorrectal (Wu et al. 2010; Wu et al. 2011). Además, se ha demostrado que los ratones *knockout* homocigotos y heterocigotos para *MUS81* tienen una predisposición a

desarrollar cáncer (McPherson et al., 2004). La correcta expresión bialélica de *MUS81* es crítica para el mantenimiento de la integridad genómica y supresión tumoral. La pérdida de un alelo de MUS81 produce acumulación de horquillas de replicación aberrantes e intermediarios de recombinación (McPherson et al., 2004). La posible asociación del gen *MUS81* con los mecanismos epigenéticos podría deberse a que la desregulación de los proceso de reparación del ADN y la acumulación del horquillas de replicación llevaría al reclutamiento y retención de las ADN metiltransferasas, ya que estas enzimas tienen una alta afinidad por sitios de ADN dañado (James et al., 2003). Así, el mantenimiento normal de la metilación durante la replicación del ADN se vería comprometido, induciendo una hipometilación global pasiva con cada ciclo de división celular (Lee et al. 2012; revisado en Wild & Flanagan 2010). Si bien todo indicaría el potencial rol de MUS81 en el desarrollo tumoral, son necesarios estudios funcionales con el fin de identificar el rol de este SNPs particular en la carcinogénesis.

En nuestro análisis detectamos una asociación positiva del genotipo CC del SNPs en *MUS81* con el nivel de metilación global del ADN, en el grupo de controles no afectados con cáncer de mama. De manera, que el genotipo CC tendría un papel protector contra la inestabilidad genómica al estar asociado a mayores niveles de metilación genómica global en sangre. Sin embargo, este resultado no es concordante al descrito por Loizidou y colaboradores (2010) quienes encontraron que mujeres portadoras del alelo C presentaban un mayor riesgo para el cáncer de mama en la población de Chipre. Estas evidencias conflictivas para la asociación podrían deberse a diferencias específicas de las poblaciones u otros factores que podrían estar interviniendo en esta asociación. El efecto protector del SNPs rs545500 *MUS81* en los controles observado en nuestro estudio, podría ser atribuido al mismo SNPs o al desequilibrio de ligamiento con otra variante génica.

Estos resultados sugieren nuevamente la importancia del efecto de las variantes genéticas en la variabilidad de la metilación global del ADN de leucocitos de los individuos, y la necesidad de integrar la información genética y epigenética de los individuos en el estudio de enfermedades complejas como el cáncer. Mientras cambios en la metilación del ADN como resultado de una enfermedad pueden ser marcadores útiles para diagnóstico temprano, variaciones en la metilación debido a influencia genética podrían ser potenciales marcadores para asignación de riesgo entre los individuos de una población.

84

#### 3.4.3 Conclusiones

En esta primera etapa del trabajo, demostramos la existencia de metilación diferencial en el ADN de leucocitos de pacientes con cáncer esporádico en comparación a un grupo control no afectado, reportándose por primera vez en un estudio caso-control en melanoma cutáneo maligno. Específicamente describimos la ocurrencia de hipometilación global del ADN en sangre periférica y una correlación negativa entre la ancestría genética africana y la metilación del ADN en leucocitos de pacientes con cáncer esporádico. A su vez, detectamos una asociación positiva del genotipo CC del SNPs en *MUS81* con el nivel de metilación global del ADN en los individuos controles no afectados con cáncer de mama. De manera, que el genotipo CC tendría un papel protector contra la inestabilidad genómica al estar asociado a mayores niveles de metilación genómica global en sangre.

Estos resultados indican un efecto del contexto genético sobre la variabilidad epigenética en nuestra población, dada por la ancestría genética de los individuos y/o variantes genéticas en genes asociados directa o indirectamente con los mecanismos epigenéticos.

Si bien no cuantificamos el riesgo a cáncer que se asocia a la hipometilación global del ADN genómico observada, estos hallazgos abren la puerta para plantear nuevos estudios para intentar ubicar en el genoma estas diferencias epigenéticas observadas en sangre mediante el estudio de perfiles de metilación del ADN a través de un abordaje genómico y sitio-específico.

# Patrones de metilación del ADN en melanoma cutáneo esporádico

## 4.1 Antecedentes

Al encontrar diferencias en la metilación global del ADN genómico entre pacientes con cáncer y controles no afectados a nivel de células normales como los leucocitos, es importante en esta etapa identificar la localización de estas diferencias de metilación en el genoma y estudiar su asociación con el cáncer. Los análisis de metilación global no proveen información sobre la posición genómica en la cual la metilación del ADN está alterada, por lo que es necesario recurrir a técnicas de mayor resolución para relacionar estos cambios epigenéticos con posibles cambios funcionales.

Tradicionalmente, los análisis de metilación del ADN eran realizados por técnicas de hibridación como *Southern blot*, las cuales permitían estudiar unos pocos sitios de restricción sensibles a la metilación dentro de secuencia de genes conocidos. Con el desarrollo de ensayos más sensibles, como la secuenciación del ADN pretratado con bisulfito y la PCR metilación-específica, se ha podido analizar en detalle múltiples sitios CpG a lo largo de secuencias de interés. Estos estudios reduccionistas han aportado importante información sobre el control local de la metilación en genes individuales (Zilberman and Henikoff, 2007). Sin embargo, hoy en día podemos sustituir los análisis a pequeña escala de genes candidatos por metodologías basadas en la detección de grandes conjuntos de marcadores anónimos seleccionados al azar, las cuales pueden proveer información valiosa a nivel tanto global como específica. Esto aumenta la probabilidad de detectar nuevos biomarcadores asociados a una patología.

#### 4.1.1 Marcadores de riesgo en melanoma cutáneo

El melanoma cutáneo es un cáncer altamente curable si es descubierto y tratado a tiempo. El pronóstico de los pacientes con melanoma varía significativamente de acuerdo al estadío tumoral. La identificación de pacientes de alto riesgo guiando la aplicación de cuidados preventivos puede significativamente reducir la mortalidad en los casos de melanoma. De esta manera, son necesarias herramientas de evaluación de riesgo efectivas que ayuden a los profesionales de atención médica primaria a identificar los individuos con alto riesgo a desarrollar un melanoma.

Los factores de riesgo para melanoma incluyen: historia familiar de la enfermedad, pigmentación clara de piel, presencia de múltiples nevos, exposición intermitente a la radiación ultravioleta y determinantes genéticos (Fang et al., 2013; Garbe and Leiter, 2009). Descontando los genes de efecto mayor como CDKN2A, se han descrito varias variantes genéticas, en genes como receptor 1 de melanocortina (MC1R), albinismo oculocutáneo tipo II (OCA2), tirosinasa (TYR); proteína 1 relacionada a tirosinasa (TYRP1), familia de transportadores de soluto 45 miembro 2 (SLC45A2); y familia de trasportadores de soluto 24 miembro 4 (SLC24A4), entre otros. Sin embargo, estas variantes comunes identificadas en estudios previos de asociación de genes candidatos (GWAS) generalmente solo afectan levemente el riesgo a melanoma (odds ratios < 1.3) (Amos et al., 2011; Barrett et al., 2012; Duffy et al., 2013; Fang et al., 2013). Por ello, todas las herramientas adicionales que puedan predecir el riesgo a cáncer en los individuos, como los marcadores epigenéticos, mejorará nuestra habilidad para identificar la población de riesgo. Los marcadores basados en la detección de MVPs en el ADN pueden ser más útiles, ya que tienen varias ventajas con respecto a otros biomarcadores candidatos como el ARN o proteínas. Es una molécula muy estable, y puede amplificarse fácilmente por técnicas moleculares convencionales. Esto permite que la tomas de muestras puedan ser más pequeñas y por ende menos invasivas para el paciente. Si bien el ARN también puede ser amplificado, es una molécula muy lábil y se limita el tiempo desde la toma de las muestras hasta la detección en el laboratorio.

# 4.1.2 Alteraciones de la metilación del ADN asociadas a pronóstico y progresión en melanoma

La transformación maligna de melanocitos sanos requiere no solo cambios genéticos estructurales sino también alteraciones epigenéticas. La detección de metilación aberrante de genes asociados a cáncer sirve como marcadores de diagnóstico temprano, evaluación de la progresión así como indicador pronóstico de respuesta a la terapia en pacientes con melanoma. Por ejemplo, niveles altos de metilación en *RASSF1A* en el tejido tumoral se correlaciona positivamente con el avance del estadío tumoral, sugiriendo que *RASSF1A* sería un buen marcador de la progresión de melanoma (Tanemura et al., 2009). Así se han descrito gran cantidad de alteraciones epigenéticas en genes involucrados en vías de señalización asociadas a crecimiento, proliferación celular y apoptosis, y metástasis (Conway et al., 2011; Li et al., 2013).

También hay un gran esfuerzo en detectar marcadores epigenéticos en ADN circulando en plasma o en células tumorales circulando en la sangre periférica, con el fin de detectar metástasis subclínica y monitorear respuesta al tratamiento (Greenberg et al., 2012; Khoja et al., 2013; Koyanagi et al., 2010). Esto ha permitido describir algunos genes metilados en las células tumorales circulantes asociados a etapas avanzadas de melanoma y a un mal pronóstico. Sin embargo, no hay estudios reportados de detección de alteración de la metilación del ADN a nivel global en genes o regiones del genoma en leucocitos de pacientes con melanoma que pudieran servir como marcadores de riesgo, y que en combinación con los factores de riesgos fenotípicos y genéticos, pudieran mejorar los actuales modelos de riesgo de melanoma cutáneo esporádico.

En este capítulo se describe la búsqueda de sitios CpG diferencialmente metilados en leucocitos de pacientes con melanoma cutáneo maligno esporádico en comparación a un grupo control. Para ello, utilizamos un diseño de dos etapas. Una primera etapa de identificación de MVPs en el ADN genómico de leucocitos de pacientes con melanoma; seguida de una segunda etapa de validación de dichas posiciones variables empleando tejidos de una cohorte independiente de pacientes con melanoma. Esta segunda cohorte, compuesta por individuos pertenecientes a población europea, nos permitirá descartar efectos intrínsecos a la población de estudio y validar la asociación entre el estado de metilación de las MVPs y la presencia de células tumorales.

## 4.2 Metodología

#### 4.2.1 Muestras de ADN utilizadas en el estudio

Se seleccionaron al azar 43 muestras de pacientes con melanoma y 43 muestras de controles no afectados de la población general de estudio (sección 2.3.2) que cumplieran los siguientes requisitos: *i*) tuvieran la medida de metilación global del ADN, y *ii*) cuya concentración de ADN fuera suficientemente alta para los análisis de AIMS, cuya reacción inicial se realiza en un volumen de únicamente 10 µl.

#### 4.2.2 Amplificación entre sitios metilados (AIMS)

En la Figura 15 se muestra un diagrama esquemático de la metodología de la amplificación entre sitios metilados descrita por Frigola y colaboradores (2002). Como se describe a continuación, utilizamos una modificación de esta técnica sustituyendo el uso de radiactividad.

Se digirió 1 µg de ADN genómico con 20 U de la enzima de restricción metilaciónsensible SmaI (New England Biolabs) durante toda la noche a 25 °C, la cual cliva dejando extremos romos (CCC/GGG). La reacción se realizó en un volumen final de 10 µl siguiendo las recomendaciones del proveedor. Luego, el ADN es nuevamente digerido con 4U del isoesquizómero insensible a la metilación Cfr9I (Fermentas) o XmaI (New England Biolabs) durante 4 horas a 37 °C, la cual deja extremos cohesivos (C/CCGGG). Se agregó a la digestión previa la mezcla de reacción completando un volumen de reacción final de 12 µl.

El adaptador fue previamente preparado mezclando en cantidades iguales los oligonucleótidos: MCF 5'-CCCGGTCAGAGCTTTGCGAAT-3' y Blue 5'-ATTCGCAAAGCTCTGA-3', ambos a una concentración de 0.4 nmol/µl, e incubando la mezcla a 65 °C durante 2 minutos, y luego dejándolos enfriar a temperatura ambiente de 30 a 60 minutos. El ADN previamente digerido fue ligado a 2 nmol de adaptador utilizando T4 ADN ligasa (New England Biolabs) durante 2 horas a 16 °C y toda la noche a 4 °C, con una posterior inactivación de la enzima a 65 °C durante 10 minutos.

Así, a la doble digestión se agregó: 5  $\mu$ l de adaptador, 0.2  $\mu$ l de T4 ADN ligasa (400U/ $\mu$ l), 2.5  $\mu$ l de buffer 10x y agua ultrapura hasta completar un volumen final de 25  $\mu$ l. De esta manera, los adaptadores se ligan a los extremos cohesivos. Los productos de ligación fueron purificados mediante el Kit AxyPrep<sup>TM</sup> PCR Cleanup (Axygen Biosciences) siguiendo las especificaciones del proveedor, removiendo así los oligonucleótidos de cadena doble o simple que quedaron sin ligarse (< 30 pares de bases). El ADN fue eluído en 25  $\mu$ l de agua y guardado a -20 °C hasta su utilización.



**Figura 15.** Diagrama de la técnica de AIMS. El ADN genómico es representado por 2 líneas paralelas, con 7 sitios de reconocimiento CCCGGG mostrados como *boxes*. Se marcan sitios no metilados (*boxes* vacíos) y metilados (*boxes* llenos con círculos indicando las citosinas metiladas). En el primer paso, la digestión de los sitios no metilados con *Smal* resulta en extremos romos, mientras los restantes sitios metilados son cortados en un segundo paso con la enzima *Xmal*, produciendo extremos cohesivos. La ligación de un adaptador a los extremos cohesivos y la amplificación por PCR resultará en la amplificación de fragmentos de ADN flanqueados por 2 sitios metilados (Jordá et al., 2009).

En una segunda etapa, los fragmentos flanqueados por dos adaptadores son amplificados por PCR utilizando diferentes sets de cebadores que hibridan a la secuencia del adaptador más el sitio de restricción y uno o más nucleótidos adicionales que se seleccionan de manera arbitraria. De manera que estos cebadores corresponden al oligonucleótido Blue extendido con la secuencia CCGGG más 1-4 nucleótidos elegidos. Se combinaron 2 cebadores diferentes en cada reacción de PCR:

## SET A BlueCTA 5'- ATTCGCAAAGCTCTGA-CCGGG-<u>CTA</u> BlueTGG 5'- ATTCGCAAAGCTCTGA-CCGGG-<u>TGG</u>

## SET B BlueCTG 5'- ATTCGCAAAGCTCTGA-CCGGG-<u>CTG</u> BlueTGG 5'- ATTCGCAAAGCTCTGA-CCGGG-<u>TGG</u>

Los cebadores extendidos con 2-4 nucleótidos arbitrarios producen un patrón de *fingerprints* AIMS de complejidad moderada, amplificando un número limitado de secuencias en un rango de 200-2.000 pb (Frigola et al., 2002).

Todas las muestras fueron analizadas en dos experimentos de AIMS independientes con los diferentes sets de cebadores (Set A y Set B). Dichas amplificaciones fueron denominadas como PCR A y PCR B, respectivamente. Las reacciones de amplificación se realizaron en el termociclador Gene Amplification PCR System 9700 (Applied Biosystems) utilizando 2 µl de ADN ligado como molde, 1.5U de Taq DNA polimerasa (Invitrogen), buffer de reacción 1X, 1.5 mM de cloruro de magnesio, 125 µM de dNTPs y 1.1 µM de cada uno de los cebadores, en un volumen final de 25 µl. Para las reacciones A y B se utilizó el mismo programa de ciclado: 30 ciclos de 15 segundos a 94 °C y 1 minuto 15 segundos a 74 °C, con una desnaturalización inicial de 3 minutos a 95 °C y una extensión final de 5 minutos a 72 °C. Los productos de PCR fueron almacenados a -20 °C hasta su utilización.

#### 4.2.3 Visualización de los resultados de AIMS

#### Electroforesis en gel de secuencia desnaturalizante

Los productos de PCR (10  $\mu$ l) fueron desnaturalizados a 98 °C durante 5 minutos. A continuación, se les adicionó 10  $\mu$ l de solución *stop* (formamida 95 %, Na<sub>2</sub>EDTA 20 Mm pH=8, azul de bromofenol 0.05 %, xilencyanol 0.05%) y se mantuvieron en hielo hasta su utilización. Las muestras desnaturalizadas (5  $\mu$ l) fueron analizadas en geles de secuencia desnaturalizantes de acrilamida 6 % con urea 8M, utilizando una cuba de electroforesis vertical CSQ33 Cleaver Scientific Ltd y peine de 72 pocillos. Las corridas electroforéticas se realizaron utilizando como buffer de corrida TBE 1x (90 mM Tris base, 90 mM ácido bórico, 2 mM EDTA pH8), con una precorrida de 30 minutos y una corrida de 4-5 h a 55W constante.

El tratamiento de los vidrios es crítico, ya que deben prepararse previamente para asegurar la adhesión del gel a uno de los vidrios que facilite la posterior tinción del gel. Los vidrios fueron limpiados con SDS 0.1%, agua ultrapura y finalmente con etanol absoluto. Luego se extendió solución de Silane (etanol 80%, ácido acético 2%, Bind Silane 0.1% en agua) en uno de los vidrios y silicona en aerosol en el otro, se dejó secar durante 15 minutos y se volvió a limpiar con etanol absoluto en una sola dirección. Una vez secos, fueron utilizados para armar el gel.

Los geles se prepararon con 50 ml de acrilamida:bisacrilamida 19:1 al 6% con urea 8M previamente preparada, 500µl de persulfato de amonio (APS) 10% y 50 µl de tetrametilen, etilen diamina (TeMed).

#### <u>Tinción con Nitrato de Plata</u>

Luego de la electroforesis, los geles fueron teñidos con nitrato de plata. Brevemente, el gel fue incubado al menos 20 minutos (o toda la noche) en etanol 10% (fijación inicial), luego 5 minutos en solución oxidante (ácido nítrico 1.54%), y se realizaba un lavado rápido (10-20 segundos) con agua destilada. Rápidamente se incubó 30 minutos en nitrato de plata 0.2%, se realizó otro lavado con agua destilada, y se incubó en solución de revelado (carbonato de sodio 2.96%, formaldehido 0.02%) hasta visualizar las bandas con claridad. Por último la reacción de revelado fue detenida incubando el gel con etanol 10% y ácido acético 0.5% (fijador final) por 10-20 minutos. Los geles, adheridos al vidrio, se secaron a temperatura ambiente y luego fueron fotografiados.

#### Interpretación de los resultados de AIMS

Las bandas de AIMS representan secuencias cortas de ADN (de hasta 2 kb de largo) flanqueadas por 2 sitios *SmaI* metilados (CCCGGG). La pérdida de metilación en alguno de los sitios impide la amplificación de la banda. Por lo tanto, para ser amplificada, una secuencia debe de cumplir 2 requerimientos: *i*) contener 2 sitios *SmaI* metilados cercanos, y *ii*) presentar homología con los nucleótidos extendidos en el extremos 3' de los cebadores. La amplificación restringida de los AIMS permite obtener *fingerprinting* de complejidad ajustable que pueda ser interpretable, con bandas en un rango de tamaño de ~200 a ~2.000 pb. Como se muestra en la Figura 15, asumiendo un apareamiento perfecto, cada nucleótido adicional en los cebadores implica un reducción de la complejidad de los perfiles obtenidos (1/5 para GC y 3/10 para AT de acuerdo con la frecuencia del nucleótido en el genoma humano) (Welsh and McClelland, 1990).

Las diferencias en la presencia o ausencia de bandas entre los perfiles de los pacientes con melanoma y los individuos controles fue determinada mediante inspección visual de los geles. No se consideró necesario realizar análisis densitométrico ya que únicamente fueron tomadas en cuenta las bandas reproducibles (en dos o más amplificaciones independientes) y claramente distinguibles del *background* del gel, y los cambios claros entre ambos grupos, considerándolos como eventos cualitativos. La pérdida de una banda representa un evento de hipometilación de CpG, mientras que la aparición de una nueva banda es indicativa de hipermetilación.

Aunque la pérdida de homocigocidad y la amplificación génica es una situación poco común debe de ser considerada en la interpretación de los resultados, ya que podrían producir cambios dramáticos en la aparición o ausencia de alguna banda.

#### 4.2.4 Caracterización de fragmentos de ADN con metilación diferencial

Las bandas diferenciales entre pacientes y controles fueron cortadas de los geles de poliacrilamida y purificadas mediante el kit QIAEX<sup>®</sup> II Gel Extraction (Qiagen) de acuerdo al protocolo recomendado por el proveedor. El fragmento aislado fue reamplificado mediante PCR con los mismos cebadores y condiciones utilizadas en el experimento de AIMS. Cuando fue necesario, el producto reamplificado fue visualizado en electroforesis de poliacrilamida desnaturalizante de tamaño pequeño y reaislada la banda repitiendo el proceso para obtener una banda única que pudiera ser secuenciada. En la Figura 16 se muestra un ejemplo de aislamiento de una banda diferencial.



**Figura 16.** Purificación de banda de interés mediante varias rondas de extracción de la banda del gel y reamplificación. A: Gel acrilamida 6%/urea 8M luego de la primera purificación y reamplificación. B: Gel acrilamida 6%/urea 8M luego de la segunda ronda de purificación y reamplificación. C: Gel agarosa 2% luego de la última purificación. La flecha indica la banda purificada. PM: marcador de peso molecular. P1/P2: pacientes con melanoma; C1/C2: controles; C-: control negativo sin ADN.

Los fragmentos de ADN aislados diferencialmente metilados fueron secuenciados de manera convencional (Sanger) en Macrogen (Korea). Las homologías de secuencia fueron analizadas utilizando los softwares BLAST (http://www.ncbi.nlm.nig.gov/BLAST) y UCSC Human Genome Browser (http://genome.ucsc.edu/).

# 4.2.5 Validación de las regiones diferencialmente metiladas identificadas por AIMS en una cohorte muestral independiente

Se analizó los niveles de metilación de sitios CpG en los fragmentos diferencialmente metilados detectados por AIMS, así como en sus cercanías, en muestras de 16 nevos benignos, 37 tumores primarios de melanoma cutáneo maligno, 36 tumores metastásicos de melanoma y 19 líneas celulares de melanoma (Tabla Suplementaria 1), utilizando la base de datos del PEBC (IDIBELL, Barcelona).

Para detectar las sondas de la plataforma HumanMethylation450 BeadChip (Illumina) que hibridan en las inmediaciones genómicas de los fragmentos de AIMS aislados se utilizó Genome Browser.

Por otro lado se realizó un análisis de expresión *in silico* de los genes asociados a los fragmentos de ADN aislados por AIMS utilizando datos públicos obtenidos de las bases de datos NCBI Gene Expression Omnibus (Edgar et al., 2002) y EBI ArrayExpress (Brazma et al., 2003). Se analizaron los datos de expresión génica de tejidos de pacientes con melanoma cutáneo obtenidos del estudio GSE-15605 (Raskin et al., 2013). Este set de datos analiza la expresión génica en 46 tumores primarios de melanoma, 12 tumores metastásicos y 16 muestras de piel normal adyacente e incluye muestras de hombres y mujeres. Se utilizó la prueba estadística Welch-t corregido por FDR para determinar las diferencias de expresión entre los 3 grupos de muestras.

#### 4.2.6 Análisis estadísticos

Las bandas variables entre los individuos resultado de AIMS fueron tratadas como sistemas binarios (presencia/ausencia) y fueron identificadas por su posición relativa en los geles de secuencia y clasificadas en los grupos A o B, dependiendo si habían sido analizadas en el PCR A o en el B, respectivamente. La presencia o ausencia de dichas bandas en el grupo de pacientes y controles fueron anotadas en una tabla de contingencia y la significancia de dichas diferencias entre ambos grupos fue estudiada mediante el programa RxC, el cual emplea el algoritmo de Metropolis para obtener una

estimación objetiva del valor p exacto cuando se trabaja con bajo número de muestras (Roff and Bentzen, 1989).

Para analizar la presencia de agrupamientos en las muestra analizadas de acuerdo a sus perfiles de metilación del ADN, realizamos análisis de agrupamientos jerárquicos utilizando distancias Manhattan y método de aglomeración completa. Únicamente se incluyeron en este análisis las bandas anónimas variables entre los individuos detectadas en cada reacción de AIMS.

La comparación de los niveles de metilación de los sitios CpG seleccionados entre las muestras de tejidos y líneas celulares se realizó mediante prueba de Wilcoxon Rank Sum, considerándose diferencias significativas con p-valor ≤0,01.

Todos los análisis estadísticos fueron realizados utilizando *R environment for statistical computing version 2.15.3* (R Development Core Team, 2013).

## 4.3 Resultados

# 4.3.1 Perfiles de metilación del ADN de pacientes con melanoma cutáneo e individuos no afectados

Con el objetivo de detectar diferencias en los perfiles de metilación del ADN y su localización en el genoma en los pacientes con melanoma cutáneo en comparación a los individuos no afectados, utilizamos una técnica de *fingerprinting* de metilación de ADN global, mediante AIMS (Frigola et al., 2002). Esta técnica determina un patrón de metilación constituido por múltiples bandas anónimas, que representan secuencias de ADN flanqueadas por dos sitios metilados, las cuales pueden ser posteriormente aisladas para la identificación de su secuencia génica.

Se analizaron muestras de ADN de leucocitos de 43 pacientes con melanoma cutáneo y 43 controles sanos mediante la amplificación con dos juegos de cebadores arbitrarios, cuyas reacciones de amplificación fueron denominadas PCR-A y PCR-B, respectivamente. Luego de analizar los patrones obtenidos con ambas reacciones de PCR, las bandas que presentaban variación en cuanto a su presencia o ausencia entre los individuos fueron enumeradas y designadas como A1, A2, B1, etc., dependiendo del set de cebadores utilizado. Únicamente se consideraron para el análisis estadístico, aquellas bandas que fueron corroboradas en dos o más amplificaciones independientes.

Algunas muestras de ADN no produjeron perfiles de AIMS consistentes y reproducibles en ninguna de las reacciones de amplificación (PCR-A y PCR-B), eliminándose del análisis. El fracaso reiterado de amplificación de una muestra fue atribuido a la baja calidad del ADN. Debido al poder de resolución variable de algunos geles, el número total de bandas informativas por muestra incluidas en los análisis estadísticos es también variable. En este estudio, consideramos únicamente cambios dramáticos en la intensidad de las bandas (presencia o ausencia), ya que probablemente reflejen mejor eventos de metilación diferencial entre pacientes y controles (Figura 17).



**Figura 17.** *Fingerprinting* de metilación obtenido mediante AIMS y visualizado en gel de poliacrilamida desnaturalizante. Las muestras fueron analizadas en triplicado, carriles: 1, 2 y 3 (muestra 1); 4, 5 y 6 (muestra 2); 7, 8 y 9 (muestra 3); 10, 11 y 12 (muestra 4); 13, 14 y 15 (muestra 5). PM: marcador de peso molecular. Se marcan en azul bandas variables entre las muestras.

Mediante la reacción de PCR-A se obtuvo perfiles claros de metilación genómica en 35 pacientes con melanoma y 35 controles no afectados, mientras que 16 muestras no amplificaron. Por otro lado, se obtuvieron perfiles de metilación mediante AIMS con PCR-B en 42 pacientes y 32 controles. Los perfiles de metilación generados por AIMS presentan aproximadamente entre 80-120 bandas informativas por caso y por reacción de amplificación.

Al comparar los perfiles de metilación de las dos reacciones de AIMS detectamos 13 bandas anónimas variables entre los individuos analizados: 9 con PCR-A, y 4 con PCR-B. Estas bandas variables observadas en los geles de AIMS corresponderían a MVPs, es decir posiciones de metilación variable entre individuos. Sin embargo, ninguna de las bandas anónimas variables se detectó exclusivamente en pacientes o en controles. En la Figura 18 se muestra como ejemplo la ampliación de un gel donde se visualizan las bandas variables (denominadas A1 y A2) entre las muestras analizadas, y se marca la presencia de tres patrones claros al comparar las muestras.



**Figura 18.** Ejemplo de perfiles de metilación obtenidos por AIMS con la reacción PCR-A en muestras de ADN de leucocitos de pacientes con melanoma y grupo control. Amplificación del gel mostrando ejemplos de los patrones de las bandas variables denominadas A1 y A2 detectados en el grupo de muestras analizadas. En letras minúsculas se indican las muestras: pacientes con melanoma (q, s, t, v, w), controles (restantes carriles). M: marcador de peso molecular.

A continuación, estudiamos si los patrones comunes de metilación de las 9 bandas variables observadas con la reacción PCR-A considerados como un conjunto podrían ser usados para clasificar las muestras en dos grupos: pacientes y controles. Sin embargo el análisis de agrupamiento no separó a los pacientes del grupo de los individuos controles (Figura 19). Esto nos llevó a analizar cada una de las bandas variables de manera independiente, intentando identificar aquellas cuya presencia o ausencia fuera diferencial entre los patrones de AIMS de leucocitos de pacientes con melanoma en comparación al grupo control.


Figura 19. Dendograma basado en agrupamiento jerárquico de los pacientes con melanoma y controles no afectados, considerando los perfiles de metilación de las 9 bandas variables obtenidas por la reacción PCR-A de AIMS. Se utilizó el método de aglomeración "completa". MM: pacientes con melanoma cutáneo maligno. MC: Controles.

De las 13 bandas con metilación variable entre los individuos analizados, únicamente 4 de ellas presentaban diferencias estadísticamente significativas en cuanto a su presencia o ausencia al comparar los perfiles de metilación obtenidos por AIMS de los pacientes con melanoma y los controles: bandas A1, A3, A4 y B2 (Tabla 6), correspondiendo a secuencias diferencialmente metiladas entre ambos grupos.

Las bandas A1, A3 y A4 fueron detectadas en mayor porcentaje en pacientes que en controles (88.6 vs 57.1%, 90 vs 48.4%, y 91.4 vs 57.7%, respectivamente) (Tabla 6), indicando eventos de hipermetilación en las secuencias de estas bandas anónimas y/o secuencias flanqueantes en los leucocitos de pacientes con melanoma. Por el contrario, la banda B2 se detectó con mayor frecuencia en el grupo de controles no afectados por cáncer (14.4 vs 46.9%), sugiriendo eventos de hipometilación en dicha secuencia en alto porcentaje de los pacientes con cáncer.

Banda	Pacientes (%)	Controles (%)	p valor
A1	88,6	57,1	0,0062
A2	82,9	74,3	0,5651
A3	90,0	48,4	0,0002
A4	91,4	57,7	0,0047
A5	27,8	33,3	0,7811
A6	72,2	57,1	0,2944
A7	80,6	80,6	1,0000
A8	41,0	33,3	0,6109
A9	25,0	20,0	0,7604
B1	97,6	100,0	1,0000
B2	14,3	46,9	0,0042
B3	100,0	96,9	0,4328
B4	100,0	96,9	0,4341

Tabla 6. Comparación de la presencia de las bandasvariables entre pacientes con melanoma y controlesno afectados. En rojo se marcan las bandasestadísticamente significativas entre ambos grupos.

## 4.3.2 Caracterización de los fragmentos de ADN con metilación diferencial entre pacientes y controles

Una vez identificados los fragmentos de ADN con metilación diferencial entre pacientes y controles, seleccionamos en una primera etapa las bandas A1 y B2 para su caracterización, mientras que las bandas A3 y A4 serán caracterizadas en una etapa posterior.

Se extrajeron ambos fragmentos (A1 y A2) de los geles de 4 perfiles correspondientes a dos pacientes con melanoma y dos controles sanos, se reamplificaron mediante PCR (Figura 16) y posteriormente fueron secuenciados. La búsqueda de homología mediante BLAST reveló la identidad de las secuencias de ADN entre sitios CpG metilados. Se identificó el sitio de restricción CCCGGG y los nucleótidos adicionados arbitrariamente en el extremo 3´ de los cebadores utilizados en la amplificación, y el tamaño de los fragmentos de ADN coincidieron con los obtenidos originalmente en los experimentos de AIMS. Las secuencias correspondientes a las bandas A1 y B2 fueron, cada una de ellas, alineadas a un único origen en el mapa del genoma humano.

El fragmento de ADN correspondiente a la banda A1 (metilada en mayor frecuencia en las muestras de pacientes con cáncer) aislada del experimento de AIMS está localizado en el cromosoma 11.q25 (posición 132115918-132116602), y mapea en el cuerpo del gen de Neurotrimina (*NTM*) (Entrez Gene ID 50863). Este fragmento, de ~790 pb, incluye un elemento SINE de 111 pb dentro de su secuencia. La secuencia del fragmento aislado tiene una media de contenido de GC~50% (Figura 20), no encontrándose dentro o cercana a una isla CpG.



**Figura 20.** Localización cromosómica del fragmento correspondiente a la banda A1 diferencialmente metilada, aislada de los perfiles de AIMS. **A)** Localización de la secuencia A1 en el contexto de las regiones génicas codificantes cercanas e islas CpG identificadas en Genome Browser. Línea vertical roja: secuencia de interés aislada. Líneas verticales verdes: islas CpGs. **B)** Ampliación del fragmento de 790 pb correspondiente a la banda A1 aislada. En barras verticales negras se indica los niveles de contenido de GC (%) a lo largo de la secuencia, y en barra horizontal gris el elemento SINE.

Por otro lado, el fragmento de ADN aislado a partir de la banda B2 (metilada en mayor frecuencia en las muestras controles) mapeó en el cromosoma 7p22.3 (posición 1597410–1597613), a ~1250 pb del extremo 5´ del gen de proteína transmembrana 184A (*TMEM184A*) (Entrez Gene ID 202915). Tampoco se encuentra dentro o cercana a islas CpG (Figura 21).



**Figura 21.** Localización de la secuencia B2 en el contexto de las regiones génicas codificantes cercanas e islas CpG identificadas en Genome Browser. Línea vertical roja: secuencia de interés aislada. Boxes verdes: islas CpGs

## 4.3.3 Análisis en tejidos y líneas celulares de las MVPs detectada en los genes NTM y TMEM184A

Para verificar si los sitios MVPs detectados por AIMS en sangre de pacientes con melanoma podrían tener una relevancia biológica, analizamos los sitios CpG en dichas regiones y en sus cercanías utilizando datos de metilación del microarreglo 450K HumanMethylation de muestras de nevos benignos, tumores primarios de melanoma, tumores metastásicos y líneas celulares de melanoma. Estas muestras de tejidos corresponden a un muestreo independiente de individuos europeos. Utilizando Genome Browser detectamos las sondas que hibridan en los fragmentos aislados A1 y B2 así como en las regiones génicas cercanas.

#### Región de banda Al (gen NTM)

En el caso de la banda A1, que mapea en el gen *NTM*, detectamos 13 sondas del microarreglo que evalúan la metilación en esta región, incluyendo la sonda cg09364022 que evalúa específicamente un sitio CpG a 37 pb del sitio SmaI-CTA del extremo 5´ del fragmento A1, y 2 sondas internas al fragmento (cg27324823 y cg019557901) (Figura 22).



**Figura 22.** Sondas evaluadas por la plataforma 450K HumanMethylation BeadChip (Illumina) dentro y en la cercanía de la secuencia de la banda A1 aislada por AIMS. Box rojo: Ubicación del fragmento A1. Figura obtenida del Genome Browser.

Comparamos entonces los valores de metilación obtenidos de la base de datos de muestras de tejidos correspondientes a la población europea de estos 13 sitios CpG entre muestras de nevos benignos, tumores primarios y tumores metastásicos de melanoma. 7 sondas presentaron metilación diferencial entre nevos benignos y tumores primarios de melanoma (prueba Wilcoxon,  $p \le 0.01$ ), incluida la sonda cercana al extremo 5<sup>-</sup> de la banda A1 (Figura 23A).

Específicamente se detecta una región hipermetilada en los tumores primarios de melanomas que se extiende 280 pb (entre cg16594910 y cg15159131) con diferencias mayores al 10% en promedio con los tejidos de nevos benignos y alta significancia estadística. Por el contrario, la región de la secuencia A1 no presentó diferencias claras entre ambos tipos de tejidos. Sin embargo, el sitio CpG correspondiente al extremo 5´ del fragmento A1 muestra una tendencia a la hipometilación en tumores primarios aunque la diferencia, siendo significativa, es menor del 10% en promedio entre ambos grupos de tejidos (Figura 23B).

Si bien la secuencia A1 mapea en la región intragénica del gen *NTM*, con el fin de evaluar el posible rol que podría cumplir este gen en el proceso tumoral decidimos analizar los niveles de expresión de ARNm utilizando los datos públicos del estudio GSE-15605, y detectamos una disminución significativa de la expresión del gen *NTM* en biopsias de metástasis de melanoma comparado con las biopsias de piel normal (p = 0.039), pero no se detecta diferencias significativas de expresión con las biopsias de tumor primario.

A		Nevo normal	T primario
cg16594910	T primario	0,003534	
Cg10554510	Metástasis	5,64E-06	0,000497
cg1719779E	T primario	0,2439	
G1/10//02	Metástasis	0,007641	0,0139
cg09214060	T primario	5,92E-04	
000214000	Metástasis	4,37E-05	0,01039
cg2/1776656	T primario	0,0009548	
Cg24770050	Metástasis	0,0001729	0,04017
cg14151947	T primario	0,0007539	
Q14151042	Metástasis	0,0001574	0,08223
cg15150121	T primario	3,47E-06	
cg12123131	Metástasis	3,50E-05	0,035
cg06217320	T primario	0,1343	
Cg00217320	Metástasis	0,1894	0,3147
cg26470719	T primario	0,4366	
Cg20470715	Metástasis	0,6446	0,5067
cg16006622	T primario	0,07945	
G10330032	Metástasis	0,8522	0,2242
cg01057010	T primario	0,002877	
Cg01007010	Metástasis	0,233	0,6889
000254022	T primario	0,01105	
Lg05564022	Metástasis	0,8368	0,1063
cg2732/1823	T primario	0,07945	
LEC/324023	Metástasis	0,6875	0,06764
cg01057001	T primario	0,6112	
CE0132/301	Metástasis	0,8522	1



Figura 23. Comparación de los niveles de metilación de sitios CpG en la región de la secuencia A1 en tejidos de nevos normales, tumores primarios de melanoma y tumores metastásicos, analizados en la plataforma Human Methylation450 BeadChip. A) Significancia estadística (p valor) de las diferencias observadas entre los distintos tejidos por prueba Wilcoxon. En rojo se marcan las diferencias significativas. B) Gráfico del promedio de nivel de metilación (valor  $\beta$ ) de cada grupo de tejidos para cada sitio CpG.

## Región de banda B2 (gen TMEM184A)

Analizando la región de la secuencia correspondiente a la banda B2, detectamos dos sondas del microarreglo (cg16443926 y cg04899361) que evalúan la metilación en el interior del fragmento aislado, y otras seis en las cercanías del extremo 5´ del gen *TEMEM184A* y su sitio de inicio de la transcripción (TSS) (Figura 24).

Scale chr7:		1.596.000	1 kb	hg19 1,598,000	1.599.000
			Your Sequence from F	Blat Search	
YourSeq			Tour bequerice if our t		
TMEM184A	<11	UCSC Genes (RefSeq,	GenBank, CCDS, Rfam,	, tRNAs & Comparati	ve Genomics)
TMEM1848	<				
		GM12878 Met	thulation 450K Bead f	Array from ENCODE/H	AIB
cg03564557		cg10633906   cg00335591	cg1644392	26	
cg02879083	r i	cg00411097	cg048993	61	
-		CQ8999986881			
		cg12583394			
		cq21368161			
		cg06194808			

**Figura 24.** Sondas evaluadas por la plataforma 450K HumanMethylation BeadChip (Illumina) dentro y en la cercanía de la secuencia de la banda B2 aislada por AIMS. Box rojo: Ubicación del fragmento B2. Figura obtenida del Genome Browser.

Comparamos los valores de metilación de los 2 sitios CpG internos al fragmento B2 y los CpG correspondientes a las sondas cercanas al TSS de *TMEM184A* entre los distintos tejidos (Figura 25). Únicamente uno de los sitios CpG del fragmento B2 presentó metilación diferencial entre nevos benignos y tumores primarios de melanoma (prueba Wilcoxon, p < 0.05) aunque la diferencia del promedio de metilación es < 10%.

La banda B2 fue detectada como región con metilación diferencial en leucocitos de pacientes con melanoma asociado a un posible evento de hipometilación en un alto porcentaje de los pacientes. También se detectó esta metilación diferencial entre tejido sano y tumoral.

Por otro lado, cuando analizamos los niveles de expresión de ARNm de *TMEM184A* utilizando los datos públicos del estudio GSE-15605, detectamos una disminución significativa de su expresión en biopsias de tumores primarios de melanoma y metástasis comparado con las biopsias de piel normal ( $p = 3.1 \times 10^{-6} \text{ y p} = 0.0089$ , respectivamente).



**Figura 25.** Comparación de los niveles de metilación de sitios CpG en la región de la secuencia B2 en tejidos de nevos normales, tumores primarios de melanoma y tumores metastásicos y líneas celulares, analizados en la plataforma HumanMethylation450 BeadChip. **A)** Significancia estadística (p valor) de las diferencias observadas entre los distintos tejidos por prueba Wilcoxon. En rojo se marcan las diferencias significativas. **B)** Gráfico del promedio de nivel de metilación (valor  $\beta$ ) de cada grupo de tejidos para cada sitio CpG.

## 4.3.4 Selección de sitios CpG candidatos para su futura validación en muestreo independiente correspondiente a nuestra población

Para establecer si los sitios diferencialmente metilados detectados por AIMS asociados a los genes *NTM* y *TMEM18A* podrían ser considerados nuevos biomarcadores epigenéticos candidatos para riesgo a melanoma cutáneo, se requiere validarlos en un set independiente de pacientes con melanoma y controles no afectados.

Con respecto al análisis de la banda A1, seleccionamos para validar primero los sitios CpG ubicados en los extremos de la región hipermetilada de 280 pb del gen *NTM* detectada en tumores primarios. De este modo, se incluyeron en el análisis la misma región estudiada en los microarreglos (sondas cg16594910 y cg15159131). Esta región, si bien cercana a la banda A1, no fue detectada por AIMS en muestras de leucocitos, por lo que decidimos incluirlas en la validación para analizar estos sitios en sangre periférica de pacientes con melanoma y grupo control, y así asegurar la validez de las comparaciones. Segundo, se seleccionaron los sitios CpG flanqueantes al sitio SmaI-CTA del fragmento A1, incluyendo el sitio correspondiente a la sonda cg09364022 a 37 pb del sitio de restricción.

Por otro lado, se seleccionaron para validar en más muestras y obtener resultados con mayor fuerza estadística, los sitios CpG internos a la secuencia B2 aislada por AIMS, ubicados hacia la posición 5' de dicha secuencia (sondas cg16443926 y cg04899361 del microarreglo). Aunque no observamos diferencias en los niveles de metilación entre nevos benignos y tumores primarios en los sitios CpG cercanos al TSS del gen *TMEM184A*, dado que la secuencia B2 se encuentra dentro de la región reguladora de este gen, decidimos ampliar el análisis evaluando los sitios CpG flanqueantes al TSS en muestras de sangre de pacientes con melanoma e individuos sanos.

## 4.4 Discusión

Una vez confirmada la presencia de diferencias en los niveles globales de metilación del ADN de leucocitos entre los pacientes con melanoma y el grupo control mediante la técnica de HPLC (descrita en el Capítulo 3), era necesario intentar detectar estas diferencias epigenéticas a nivel de secuencia.

En este trabajo, por primera vez se describe la búsqueda y caracterización de regiones con metilación diferencial en el ADN genómico de leucocitos de pacientes con melanoma cutáneo maligno, mediante la aplicación de una técnica sencilla y de bajo costo como AIMS. De esta forma, pudimos detectar variabilidad en los patrones de metilación del ADN en sangre periférica entre los individuos, y caracterizar dos regiones genómicas con metilación diferencial en los pacientes con melanoma en comparación al grupo control no afectado.

#### 4.4.1 Aplicación de la técnica de AIMS en estudios poblacionales

Implementamos en el laboratorio una modificación de la técnica de AIMS, que permite analizar perfiles de metilación del ADN genómico de manera sencilla y en un gran número de muestras, detectando al mismo tiempo eventos de hipermetilación e hipometilación del ADN.

Hasta hoy en día, las aplicaciones de la técnica de AIMS para investigar los cambios epigenéticos en cáncer incluyen la identificación de hipermetilación recurrente asociada con silenciamiento génico (Frigola et al., 2005a, 2006); la búsqueda tanto de hipometilación como hipermetilación en líneas celulares tumorales con una función de la metilación del ADN alterada (Paz et al., 2003; Sadikovic et al., 2004); y estimación del grado de metilación del ADN alterado en cáncer (Frigola et al., 2005b; Rodriguez et al., 2006). Esta técnica también ha sido aplicada para identificar diferencias epigenéticas surgidas durante la vida de mellizos monocigóticos (Fraga et al., 2005). En todos los casos, los estudios se realizaron sobre muestras pareadas de tejidos correspondientes a dos situaciones a ser evaluadas (por ejemplo: tejido normal/tejido sano del mismo individuo, par de mellizos).

De acuerdo a nuestro conocimiento, nuestro trabajo describe por primera vez la aplicación de la técnica de AIMS en un estudio de asociación a nivel poblacional, donde no se comparan muestras de tejidos pareados, sino dos grupos de muestras correspondientes a pacientes y controles. De esta manera hemos distinguido variación interindividual de la metilación del ADN genómico de leucocitos de los individuos analizados.

Hay que tener en cuenta que la utilización de enzimas de restricción metilaciónsensibles para analizar metilación implica que no se analizan todos los sitios CpG del genoma, sino solo los CpG en secuencias CCCGGG, lo que corresponde a menos del 8% del total de sitios CpG genómico. A su vez, al utilizar cebadores arbitrarios con adición de nucleótidos en su extremo 3', se disminuye el número de secuencias amplificadas reduciendo aún más el número de sitios CpG analizados; y se pierde sensibilidad en el proceso de tinción de los geles de secuencia.

A pesar de esta reducción de la información, los sitos analizados se encuentran distribuidos en todos los cromosomas ya que la distribución de los sitios CCCGGG a lo largo del genoma humano es azarosa, aunque el genoma mitocondrial carece de sitios CCCGGG (Tanas et al., 2010). Al reducir la complejidad de la información obtenida, permite ser una técnica fácil de implementar en el laboratorio y de rápido análisis, sin perder la representación de todo el genoma nuclear y con suficiente información como para evaluar la existencia de metilación variable en un gran número de secuencias. Esto hace que la relación costo-beneficio de la técnica sea muy favorable para evaluar gran número de muestras en simultáneo en estudios epidemiológicos.

Un análisis cualitativos de la representación genómica de las secuencias posibles de ser evaluadas por AIMS reveló su enriquecimiento con secuencias repetidas, en su mayoría entre 100–1000 pb, que incluyen varias clases de repetidos en tándem y dispersos, como Alu, LTR, repetidos periteloméricos, entre otros (Tanas et al., 2010). Por otro lado, la representación de islas CpG es lo suficientemente frecuente ya que alrededor del 70-80% de las islas CpG contienen al menos dos sitios de restricción *SmaI* separados por  $\leq$  1 kb (Huang et al., 2010), lo que hace que esta metodología permita estudiar también variabilidad de metilación en islas CpG.

En este trabajo buscábamos detectar procesos de metilación sistémicos que puedan afectar a todo el organismo o varios tejidos, al ser heredados o surgir en etapas embrionarias y/o posteriores al nacimiento. Los efectos podrían dar lugar a cambios en regiones o genes que aunque no estén directamente involucrados con el desarrollo de la enfermedad son marcadores diferenciales.

De esta manera, analizamos por AIMS los perfiles de metilación en el ADN de leucocitos de 43 pacientes con melanoma y 43 controles. La primera conclusión que podemos sacar de los resultados crudos es que los perfiles de metilación del ADN genómico de leucocitos es variable entre los individuos, detectándose 13 bandas polimórficas. La variabilidad correspondiente a la presencia o ausencia de estas bandas no agrupan a pacientes y controles en análisis de agrupamientos jerárquicos, lo que indica que la variabilidad observada en los perfiles no se debe únicamente a la condición de la patología (paciente/control). Estos resultados van en concordancia con otros estudios epigenómicos poblacionales que han demostrado que los patrones de metilación del ADN varían entre los individuos y las regiones genómicas (Johnson and Tricker, 2010). En este contexto, existen fuertes evidencias del rol directo y la relevancia de la epigenética sobre la variabilidad fenotípica humana (Czyz et al., 2012). Actualmente es un desafío determinar la casualidad y la contribución relativa del medioambiente, los factores genéticos, y los factores estocásticos sobre la variabilidad epigenética. Fraga y colaboradores (Fraga et al., 2005) proponen que el epigenoma es heredado en gran medida al nacimiento pero epimutaciones ocurren y se acumulan a lo largo de la vida, y sus orígenes surgen como resultado de una combinación de factores externos y una deriva epigenética produciendo cambios en la metilación. Por otro lado, Kaminsky (2009) compara la variación epigenética inducida estocásticamente con la inducida por el medioambiente, considerando los eventos estocásticos como la explicación más importante en la discordancia fenotípica de mellizos monocigóticos. Recientemente, también se ha demostrado la existencia de variación epigenética natural en poblaciones vegetales, donde la mayoría de las MVPs no estarían ligadas al genotipo (Schmitz et al., 2013).

Si bien la gran mayoría de las bandas polimórficas detectadas en el análisis de AIMS corresponden a variación natural en los patrones de metilación entre los individuos, 4 de las bandas detectadas son diferenciales entre pacientes con melanoma y controles, indicando en este caso su posible asociación con el cáncer. De manera interesante, al igual de lo que sucede en el tejido tumoral, detectamos eventos de hipermetilación así como hipometilación del ADN en leucocitos de pacientes en comparación a los controles. Esto confirma la utilidad de la técnica de AIMS tanto para el estudio de variabilidad interindividual en poblaciones, así como para estudios epidemiológicos en cáncer permitiendo detección de MVPs en células sanguíneas de pacientes con cáncer que pudieran estar asociadas al desarrollo tumoral.

Es importante resaltar que, dado que no tenemos los datos de los estadios tumorales ni el seguimiento clínico luego de la toma de la muestra de los pacientes con melanoma, no podemos descartar que la variabilidad epigenética observada en células sanguíneas de los pacientes se deba en parte a la presencia de células tumorales circulando en la sangre en algunos pacientes. En pacientes con melanoma se han detectado células tumorales viables circulando en sangre (Vereecken et al., 2012). Establecer si los eventos de metilación diferencial del ADN de leucocitos en los pacientes con melanoma se produjeron *a priori* del desarrollo tumoral o como consecuencia de la patología requerirá otros estudios. Aun no sabiendo su origen, la detección de estas posiciones de metilación variable tienen un gran potencial como marcadores de riesgo (si son detectados en individuos asintomáticos) o diagnóstico temprano de la enfermedad.

Si bien en este trabajo se incluyeron únicamente los resultados obtenidos en la detección de secuencias con metilación diferencial en leucocitos de pacientes con melanoma, en nuestro laboratorio se está llevando a cabo el mismo abordaje para cáncer de mama. En este caso, se detectaron por AIMS 22 bandas polimórficas entre los individuos analizados, siendo dos de ellas diferencialmente metiladas entre pacientes con cáncer de mama esporádico y su grupo control (p < 0.01), ambas correspondientes a eventos de hipometilación en pacientes con cáncer (Da Silveira, 2012). Este trabajo se encuentra en etapas iniciales, por lo que los resultados deben ser corroborados, e comenzar la caracterización de estas regiones. Pero nuevamente demuestra el uso potencial de esta técnica en la detección de posibles biomarcadores de cáncer mediante estudios caso-control.

## 4.4.2 Regiones con metilación diferencial del ADN de leucocitos en pacientes con melanoma cutáneo

En este estudio, identificamos cuatro regiones con metilación diferencial en el ADN de leucocitos de pacientes con melanoma cutáneo, caracterizando dos de ellas y

validando los resultados en un grupo independiente de muestras de tejidos de pacientes con melanoma de origen europeo. La detección de patrones de metilación alterados en sangre de los pacientes sugiere su potencial uso como marcador de riesgo o diagnóstico temprano de melanoma cutáneo esporádico.

Los estudios de búsqueda de marcadores epigenéticos asociados a melanoma realizados hasta hoy en día en sangre periférica se basan en la detección de alteraciones en la metilación de ADN libre o de ADN de células tumorales circulantes. En estos estudios, el abordaje de genes candidatos es más frecuente que los análisis genómicos, posiblemente debido a la cantidad limitada de ADN disponible. Sin embargo, mediante un abordaje genómico en el ADN de leucocitos, varios marcadores de metilación potenciales fueron identificados en distintos tipos de cáncer, incluyendo ovario, páncreas, vejiga, mama y de pulmón (Li et al., 2012). En melanoma cutáneo no hay reportes de marcadores epigenéticos potenciales detectados en el ADN de leucocitos de pacientes.

Mediante análisis genómico detectamos dos secuencias de ADN con metilación diferencial en uno o ambos de sus extremos en leucocitos de pacientes con melanoma. Uno de los fragmento de ~790 pb mapeó en la región intragénica del gen Neurotrimina (*NTM*), mientras que el segundo fragmento mapeó a ~1.250 pb del extremo 5´ del gen *TMEM184A*.

#### <u>Neurotrimina</u>

El gen *NTM* codifica una proteína de 344 aminoácidos, de tipo IgLON, una subfamilia de inmunoglobulina de moléculas de adhesión celular ancladas a glicosilfosfatidilinositol. Esta familia comprende la proteína de unión a opioide (*OPCML*), la proteína de membrana asociada al sistema límbico (*LSAMP*), neurotrimina (*NTM*) y el regulador del crecimiento neuronal 1 (*NEGR1*). Se sugiere que las IgLON podrían tener un importante rol en la adhesión celular y reconocimiento célula-célula, mediante agrupamiento de proteínas en la superficie celular (McNamee et al., 2002). *OPCML* comparte la homología más alta con *NTM* entre los cuatro miembros de la familia IgLON, compartiendo 92% de identidad en el primer dominio Ig, y en el segundo y tercer dominio Ig 70% y 66% de identidad, respectivamente (Cui et al., 2008). A su vez, los genes *NTM* y *OPCML* se localizan muy cercanos en la misma

112

región cromosómica en orientación opuesta, donde el extremo 5´ del gen *OPCML* se encuentra a solo 80 kb del extremo 3´ del gen *NTM*. Se han detectado una gran cantidad de transcriptos alternativos *NTM* en distintas líneas celulares (UCSC Genome Browser).

Los estudios de la proteína NTM se centran sobre el sistema nervioso central, donde tiene una alta expresión y ha sido asociada a una posible función en el desarrollo neuronal, crecimiento del axón y función sináptica. La región 11q25, donde se ubica el gen de neurotrimina está localizada en una región ligada a la enfermedad de Alzheimer, desorden de déficit de atención, autismo y depresión, y específicamente SNPs en el gen *NTM* fueron asociados al coeficiente intelectual humano en estudios familiares (Pan et al., 2011). La expresión y función de la neurotrimina en tejidos periféricos normales no está caracterizada, habiendo sido estudiada únicamente en útero donde está regulada por estrógeno. Esto sugiere la idea que *NTM* participa en múltiples procesos que gobiernan dinámicamente la integridad de los blancos de inervación simpática periférica (Krizsanagbas et al., 2009).

Por otro lado, se describió la expresión de IgLONs en células de ovario y carcinoma renal, y su expresión está alterada en el cáncer, sugiriendo que las IgLONs podrían actuar como genes supresores de tumores (Chen et al., 2003; Ntougkos et al., 2005). Probablemente actúen como supresores de tumores mediante la interacción entre IgLON distintas formando heterodímeros involucrados en la transducción de señales. Específicamente, la expresión de todas las IgLON difiere significativamente entre ovarios normales y cáncer de ovario epitelial a nivel del ARN. Los niveles de expresión de OPCML, LSAMP y NEGR1 están reducidos en cáncer de ovario, mientras que los de NTM se encuentran aumentados (Ntougkos et al., 2005). Sin embargo, al analizar los niveles de expresión de NTM en las bases de datos públicas, se detectó una disminución significativa en los niveles de ARNm de este gen en metástasis de melanoma en comparación a piel normal (Raskin et al., 2013), sin dejar claro el papel de la NTM en la tumorigénesis. Existen suficiente evidencia de que los tumores sólidos y menos invasivos generalmente expresan altos niveles de proteínas de adhesión celular en comparación a los tumores altamente invasivos y diseminados como melanoma (Katto and Mahlknecht, 2011). Dado que el cambio significativo de expresión génica de NTM sólo fue evidenciado en melanoma metastásico y no en el tumor primario, y que este gen codifica una proteína de adhesión, podríamos sugerir que NTM podría jugar un rol en la alta capacidad de metástasis de melanoma. NTM tendría un papel importante en la internalización de hedgehog, y recientemente se describió el involucramiento de la vía de señalización de hedgehog en la progresión de melanoma (O'Reilly et al., 2013).

La causa de la reducción en expresión de *OPCML* descrita en cáncer de ovario epitelial, carcinoma de pulmón, carcinoma nasofaríngeo, cáncer de esófago, así como en múltiples líneas celulares de carcinomas es fundamentalmente epigenética, lo que hace suponer un control epigenético alterado de estos genes durante el desarrollo tumoral (Cui et al., 2008; Sellar et al., 2003), e indica que la metilación de estos genes podría ser un biomarcador epigenético para varios tipos de cáncer.

Aunque se han reportado 13 mutaciones somáticas de cambio de sentido a lo largo del gen *NTM* en tumores de melanoma cutáneo maligno (en comparación a tejido normal pareado de los mismos individuos) (The Cancer Genoma Atlas, 2013), la importancia de estas mutaciones y la funcionalidad de este gen en la piel no ha sido descrita.

Nosotros detectamos variación en la metilación del ADN de leucocitos de pacientes con melanoma en el intrón 4 del gen *NTM*. Al validar los resultados de AIMS en tejidos normales y tumorales, ampliando la región de análisis, detectamos también una región cercana hipermetilada de 280 pb dentro del mismo intrón en los tumores primarios de melanoma. Las consecuencias de esta variación en la metilación en secuencias intragénicas necesitan ser elucidadas, sin embargo en los últimos años ha adquirido gran importancia los eventos de metilación intragénica como mecanismos de modulación de la expresión génica ya sea mediante la regulación de promotores alternativos, potenciadores, expresión de ARNm no codificantes intragénicos, así como regulando los sitos de splincing (Kulis et al., 2013; Shenker and Flanagan, 2012).

Además, tenemos que tener en cuenta que la secuencia detectada por AIMS en el gen de *NTM* incluye un elemento SINE en su interior. Los SINEs son una clase de retrotransposones que amplifican su número de copias en el genoma huésped por retrotransposición. Más de un millón de copias de SINEs están presentes en el genoma de mamíferos, contribuyendo al 10% del total de secuencias genómicas. Los elementos SINEs se transcriben por la ARNpol III y su transcripción está regulada por mecanismos epigenéticos como la metilación del ADN. Estudios recientes revelaron que estos elementos juegan roles funcionales, por ejemplo como potenciadores distales, bordes de eucromatina y heterocromatina, reposicionamiento de nucleosomas y sitios de unión a factores de transcripción (Ichiyanagi, 2013). Por lo que podría existir otros

mecanismos involucrados en la asociación de esta región diferencialmente metilada con melanoma cutáneo.

La detección de metilación diferencial en un elemento repetitivo intragénico en el gen *NTM* es consistente con estudios previos en los cuales la mayoría de los sitios CpG asociados a cáncer fueron encontrados en secuencias intragénicas (Shenker and Flanagan, 2012). Incluso, se ha demostrado que existe mayor variabilidad en la metilación del ADN dentro de los cuerpos génicos que dentro de promotores (Flanagan et al., 2009). Así, estos datos plantean la posibilidad que la susceptibilidad al cáncer podría ser influenciada tanto por la metilación más variable en elementos repetitivos intragénicos como por la metilación menos variable en islas CpG.

#### <u>TMEM184A</u>

La segunda secuencia diferencialmente metilada en leucocitos de pacientes con melanoma y controles, y luego validada en tejidos mapea en la región promotora del gen *TMEM184A*, que codifica una proteína transmembrana de 449 aminoácidos. Esta proteína, estudiada únicamente en ratón, tiene función durante la determinación sexual en el desarrollo embrionario, y también se ha detectado su expresión en tejidos no gonadales como estómago, intestino, vejiga, próstata, páncreas, glándulas salivales y glándulas mamarias (Best and Adams, 2009). Aunque se sugirió una actividad quinasa para *TMEM184A*, su secuencia de aminoácidos no alinea con ningún dominio conocido, por lo que su función es aún desconocida (Svingen et al., 2007).

Dado que esta proteína no ha sido caracterizada en humanos y muy poco se conoce sobre su función en los distintos tejidos y su mecanismo de acción o interacción con otras proteínas, es difícil hipotizar sobre el papel que podría cumplir durante el desarrollo tumoral. Sin embargo, en estudios genómicos de expresión de ARNm publicados en bases de datos disponibles, se detectó una disminución significativa de la expresión del gen *TMEM184A* en biopsias de tumores primarios y metástasis de melanoma cutáneo en comparación a piel normal (Raskin et al., 2013), lo que podría abrirnos una pequeña ventana para seguir investigando este posible nuevo candidato. Así, de ser validado en un muestreo mayor e independiente, será necesario realizar estudios funcionales en líneas celulares para poder elucidar la implicancia de los cambios de expresión de este gen, así como las consecuencias de la variación de la metilación en su región promotora.

#### Siguiente paso en el estudio de NTM y TMEM184A como biomarcadores

Mediante la técnica de AIMS, identificamos dos regiones en los genes NTM y TMEM184A, con metilación diferencial significativa en el ADN de leucocitos de pacientes con melanoma cutáneo. Por sí solo, estos resultados no son suficientes para considerar estas regiones como biomarcadores de riesgo a melanoma, ya que ninguna de las dos bandas detectadas en los perfiles de AIMS fueron exclusivas de los pacientes o de los controles, existiendo un solapamiento entre ambos grupos. Además, como la mayoría de los abordajes de tamizaje (screening), los resultados de AIMS requieren confirmación mediante el análisis con otra técnica en un mayor número de muestras. Con este fin, estamos implementando en nuestro laboratorio la cuantificación de metilación del ADN mediante MS-HRM de las dos regiones de interés y sus cercanías. Dado que los resultados de AIMS obtenidos son cualitativos (presencia o ausencia de la banda) para cada individuo, con la validación por MS-HRM no solo obtendremos datos en un muestreo independiente y con mayor poder estadístico, sino también datos de carácter cuantitativo para cada individuo. Esto nos permitirá realizar análisis de curvas ROC para cada uno de los marcadores candidatos, y así determinar su poder de clasificación de pacientes y controles. A su vez, es posible aumentar la sensibilidad en combinación con otros biomarcadores de metilación del ADN y/o factores de riesgos epidemiológicos o genéticos, si estos marcadores proveen un efecto aditivo (Marsit et al., 2011; Pedersen et al., 2011).

## 4.4.3 Conclusiones

En este trabajo por primera vez se identifica metilación sitio-específica en ADN de leucocitos asociada a pacientes con melanoma cutáneo maligno. Mediante una técnica de *screening* de un porcentaje del genoma, detectamos cuatro MVPs diferencialmente metiladas en leucocitos de pacientes con melanoma cutáneo en comparación a individuos sanos. Dos de estas regiones, fueron caracterizadas y validadas en muestras

de tejidos de individuos con melanoma cutáneo de origen europeo. Así seleccionamos los genes *NTM* y *TMEM184A*, ambos codificadores de proteínas transmembranas, para su validación en un muestreo mayor de nuestra población. Las restantes dos MVPs serán en un futuro caracterizadas siguiendo el mismo diseño experimental.

Finalmente, estos resultados también sugieren la necesidad de analizar epigenomas completos, con enfoques no sesgados, para poder estudiar el complemento completo de la variabilidad de la metilación del ADN a lo largo del genoma, y así poder detectar nuevas regiones asociadas a cáncer.

El alcance sin precedente del estudio de todo el genoma y la precisión de nucleótido con lo que ahora podemos mapear epigenomas humanos fue posible gracias al avance de las tecnologías de alta performance, como los microarreglos y la secuenciación masiva (NGS). Por lejos, la mayor ventaja de NGS es su capacidad de analizar todo el genoma de manera imparcial y exhaustiva. Sin embargo, esta técnica requiere un alto nivel de especialización y son aún caras y consumen mucho tiempo (Rivera and Ren, 2013). Por eso, la nueva plataforma de microarreglos de Illumina de más de 450.000 sitios CpG para estudios de metilación es una herramienta promisoria.

De esta manera, luego de haber detectado diferencias en los niveles de metilación global del ADN de leucocitos de pacientes con cáncer, y haber detectados de manera específica variaciones a nivel de secuencia con una estrategia de análisis del genoma parcial, en el siguiente capítulo abordamos de manera más amplia el estudio de sitos CpG diferencialmente metilados en pacientes con cáncer mediante microarreglos de metilación.

# Metilación del ADN sitioespecífica en cáncer de mama esporádico

## **5.1 Antecedentes**

La mayoría de los estudios de metilación del ADN como biomarcador potencial de cáncer utilizan tejido tumoral. Sin embargo, un número creciente de estudios han usado fluidos corporales como orina, leche materna, plasma, suero y sangre periférica (Li et al., 2012). Hasta hace 5 años eran muy pocos los estudios publicados donde se investiga la asociación de metilación loci-específica en leucocitos y riesgo a cáncer, por lo que los datos son aún limitados y deben ser validados en cada población (Terry et al., 2011).

Se reportó por primera vez metilación diferencial loci-específica en el ADN de leucocitos de sangre periférica en cáncer de pulmón (Woodson et al., 2001). En este estudio caso-control (n=100), los investigadores sugirieron que el estado de metilación en el ADN extraído de sangre periférica reflejaría el estado del ADN del tejido pulmonar. Ellos identificaron una correlación entre hipometilación del gen p53 en el ADN de sangre y cáncer de pulmón. En cáncer de mama, la metilación de loci específicos en genes diana de ER $\alpha$  y genes diana del grupo Polycomb de ADN de células de sangre periférica fue asociada con riesgo a cáncer de mama (Widschwendter et al., 2008). Flanagan y colaboradores (2009) compararon el patrón de metilación del ADN de sangre periférica utilizando microarreglos de metilación de diseño cubriendo 4 Mb con 51 genes candidatos, en 14 individuos con cáncer de mama bilateral y 14 controles sanos. Los resultados validaron sus hallazgos iniciales en el gen *ATM* en 190 pares de casos y controles, así como probaron la hipótesis de que algunos cambios epigenéticos sistémicos podrían ser detectados en el ADN de sangre periférica en cáncer de mama. Años después el mismo grupo de investigadores utilizaron muestras

prediagnóstico de pacientes con cáncer de mama, excluyendo la posibilidad de que la variabilidad en la metilación en el ADN de sangre estuviera influenciada por la presencia de células tumorales o tratamiento en los pacientes (Brennan et al., 2012). Describieron una asociación de la metilación del gen *ATM* en sangre de pacientes tomada hasta 11 años antes del diagnóstico. Estos datos sugieren que esta metilación diferencial en el ADN de células sanguíneas representa un marcador estable de predisposición más que un evento tumorigénico temprano.

Utilizando un abordaje de análisis de todo el genoma usando ADN de leucocitos, se han identificado en los últimos años potenciales biomarcadores de metilación en varios tipos de cáncer. La publicación del primer artículo donde se utiliza la plataforma de microarreglos Illumina Infinium HumanMethylation27 (27.578 sitios CpGs) para diferenciar casos de cáncer de ovario de un grupo control sugiere el potencial de este abordaje como herramienta para detectar o predecir ciertos tipos de cáncer (Teschendorff et al., 2009). Posteriormente esto fue corroborado en un estudio de cáncer de vejiga donde demuestran, con alta exactitud, la capacidad de distinguir individuos con cáncer de vejiga de los controles utilizando un modelo que considera el patrón de metilación en sangre de 9 loci, edad, sexo, y estatus de fumador del individuo (Marsit et al., 2011).

Con la validación de la nueva plataforma de microarreglos de 485.577 sitios CpG para estudios de metilación (Illumina Infinium HumanMethylation450 BeadChip) se amplían las posibilidades de detección de nuevos marcadores epigenéticos asociados a enfermedades (Sandoval et al., 2011).

El objetivo del trabajo descrito en este capítulo es identificar nuevos marcadores epigenéticos específicos de cáncer de mama en sangre en la población uruguaya. Para ellos se analizó el perfil de metilación utilizando ADN de leucocitos de pacientes con cáncer de mama esporádico y controles sanos, mediante la tecnología Infinium DNA Methylation BeadChip que cubre más de 480.000 sitios CpG a lo largo de todo el genoma.

## 5.2 Metodología

#### 5.2.1 Descripción de la población de estudio

Se seleccionaron 24 muestras de ADN de leucocitos de sangre periférica de pacientes con cáncer de mama y 11 muestras de controles no afectados de la población general de estudio (sección 2.3.2) que cumplieran los siguientes requisitos: *i*) tuvieran la medida de metilación global del ADN, *ii*) tuvieran datos de ancestría genética individual, *iii*) pasaran los controles de calidad en cuanto a integridad, cantidad y pureza necesaria para el análisis de microarreglos, y *iv*) los individuos controles no fueran fumadores. Se seleccionaron los pacientes y controles de tal manera que no hubiera diferencias en cuanto a la edad y ancestría entre los grupos, y que la mayoría de los pacientes seleccionados correspondiese a estadios tumorales iniciales (I y II). Las mujeres con cáncer de mama fueron codificadas como KP, mientras que las mujeres controles como KC.

#### 5.2.2 Infinium HumanMethylation450 BeadChip

Analizamos simultáneamente el estatus de metilación de 485.577 sitios CpG mediante tecnología Infinium, utilizando el microarreglo HumanMethylation450K BeadChip (Illumina, Inc. Utilizando este panel se obtiene información cuantitativa de la metilación de los sitios CpG con una resolución de un único nucleótido, cubriendo 361.766 CpGs en genes codificadores de ARNm, 4.168 CpGs en ARNs no codificantes (cuya mayoría corresponden a microARNs), y 119.830 CpGs no asociadas a transcriptos anotados. Los ensayos de microarreglos fueron realizados en el Laboratorio de Epigenética del cáncer, Programa de Epigenética y Biología del Cáncer (PEBC, www.pebc.cat), Instituto de Investigaciones Biomédicas de Bellvitge (Barcelona, España) bajo la supervisión de la Dra. María Berdasco.

Todas las muestras fueron analizadas para su integridad, cantidad y pureza mediante electroforesis en gel de agarosa, cuantificación por fluorescencia utilizando Picogreen® y por absorbancia en NanoDrop. Las muestras fueron distribuidas al azar en

placas de 96 pocillos. Se realizó la conversión con bisulfito de 500 ng de ADN genómico utilizando el EZ DNA Methylation kit (ZymoResearch) de acuerdo a las indicaciones del proveedor. Las muestras de ADN convertido fueron amplificadas mediante Whole Genome Amplification (Quiagen) seguido de una fragmentación enzimática, precipitación y resuspensión. Las muestras resuspendidas fueron hibridadas en HumanMethylation450 BeadChip (Illumina) por 16 h a 48 °C, y luego lavadas. Se realizó una extensión de nucleótido simple con dideoxinucleótidos marcados. Los datos de los experimentos de hibridación del microarreglo fueron procesados utilizando el paquete lumi (Du et al., 2008) disponible para Bioconductor (Gentleman et al., 2004) en el lenguaje estadístico R. Se realizaron ajustes de balance de color y normalización quantil con el fin de normalizar las muestras entre los dos canales de color. El nivel de metilación del ADN se muestra como valores  $\beta$  (rango entre 0-1) y fueron calculados como el cociente entre la señal metilada y la suma de las señales metiladas y no metiladas más 100. Se determinó el p valor de detección para cada valor  $\beta$ , el cual representa el poder estadístico para la diferencia entre la señal de cada sonda con el *background* (el promedio para todos los controles negativos). Los valores  $\beta$  con un pvalor de detección > 0.01 fueron removidos de posteriores análisis. A su vez, se eliminaron del análisis 1947 CpGs que tenían un p-valor >0.001 en más del 6% de las muestras totales, y 23631 CpGs que presentaban SNP de una frecuencia > 1% (1000 Genomes Project Consortium 2010) en la secuencia de la sonda. Dado que todas las muestras de ADN pertenecían a mujeres, no se eliminaron las sondas del cromosoma X.

## 5.2.3 Identificación de CpG diferencialmente metilados entre pacientes y controles

Todos los análisis estadísticos fueron realizados utilizando el lenguaje R versión 2.15.3 (R Development Core Team, 2013) y análisis no paramétricos. Todos los p-valores presentados fueron corregidos por FDR (Tasa de falso descubrimiento) al menos que se aclare lo contrario.

Para definir sitios CpG diferencialmente metilados entre pacientes con cáncer de mama y controles, excluimos las muestras "outliers" con un perfil de metilación genómica atípico, y comparamos cada uno de los CpGs del microarreglos entre pacientes con cáncer de mama y controles no afectados aplicando la prueba de Wilcoxon Rank sum para los valores  $\beta$  normalizados de cada grupo. Los p-valores fueron ajustados utilizando FDR, y los CpGs con p-valor corregidos <0.05 fueron seleccionados y denominados CpG diferencialmente metilados (CpGDM).

Las muestras fueron organizadas por agrupamiento jerárquico utilizando el método de aglomeración "completa" para distancias euclidianas, utilizando los CpG diferencialmente metilados detectados. La fuerza estadística del agrupamiento fue evaluada mediante *bootstrap* de 10.000 remuestreos usando el paquete pvclust (Suzuki and Shimodaira, 2006) disponible en el lenguaje estadístico R. También calculamos 10.000 agrupamientos usando 45.000 CpGs al azar para asegurar que el agrupamiento no fue generado basado en diferencias genómicas entre las muestras.

Por otro lado, con el fin de explorar la similitud de los datos de metilación entre el grupo de pacientes y los controles no afectados, se realizó un análisis de escalamiento multidimensional (MDS) tomando como variables los valores de metilación (valor  $\beta$ ) de los CpG diferenciales detectados de cada individuo. De esta manera se representan de manera parsimoniosa las distancias euclidianas entre los individuos en dos dimensiones.

Se evaluó el efecto de potenciales confusores en nuestro análisis al determinar las CpGDM mediante modelo de regresión lineal generalizada. Para ello se consideró como variables independientes la edad, IMC y estatus de fumadora. La edad se analizó como variable continua, mientras que las restantes como variables cualitativas: se clasificó el IMC en 3 rangos: < 25; 25-29,9; >30 y el estatus de fumadora como si/no considerando al individuo como fumador tanto si lo era en el momento de la toma de la muestra como en el pasado (Tabla Suplementaria 4).

## 5.2.4 Anotación génica, análisis de ontología génica y enriquecimiento de vías moleculares

Los sitios CpG diferencialmente metilados, cuando fue posible, se asignaron a su correspondiente locus génico, usando la construcción HG19 del genoma humano (Human Feb. 2009 (GRCh37/hg19)), así como a los distintos compartimentos genómicos: promotores ( $\pm$  2 kb alrededor del sitio de inicio de la transcripción), exones, intrones y regiones intergénicas. La asociación de dichos genes a cáncer fue analizada

mediante análisis de anotación génica en la base de datos Genes to Systems Breast Cancer Database (G2SBC, www.itb.cnr.it/breastcancer/) (Mosca et al., 2010) y búsqueda en la base de datos de Cancer Gene Census (COSMIC, http://cancer.sanger.ac.uk/cancergenome/projects/census/) (Forbes et al., 2011). Se utilizó Database for Annotation, Visualization, and Integrated Discovery (DAVID v6.7) (Huang et al., 2009) y Kyoto Encyclopedia of Genes and Genomes (KEGG database) para identificar enriquecimiento de vías moleculares y términos de ontología génica. Se usó los parámetros establecido por defecto, y un p-valor < 0.05 sin corregir fue considerado un enriquecimiento significativo.

Por otro lado, se analizó el enriquecimiento de las CpGDM en cuanto a su localización genómica y contexto isla CpG en promotores. Para ello se comparó mediante prueba de Fisher el número de CpGDM en cada categoría con lo esperado de acuerdo a la representación de la totalidad de los sitios CpG en el microarreglo 450K HumanMethylation BeadChip utilizado.

## 5.2.5 Análisis de bases de datos de expresión génica

Se realizó un análisis de expresión génica *in silico* de los CpGDM utilizando datos públicos obtenidos de las bases de datos NCBI Gene Expression Omnibus (Edgar et al., 2002) y EBI ArrayExpress (Brazma et al., 2003). Se analizó los datos de expresión génica de los genes asociados a los CpGDM en sangre de pacientes con cáncer de mama y grupo control con mamografías normales obtenidos del estudio GSE-27562 (LaBreche et al., 2011). Este set de datos analiza la expresión génica en 57 mujeres con cáncer de mama, 37 mujeres con tumor benigno de mama y 31 mujeres con mamografía normal, y se utilizó la prueba estadística Welch-t corregido por FDR para determinar las diferencias de expresión entre los tres grupos. Por otro lado se realizó un meta-análisis en la plataforma NextBio Research® (Kupershmidt et al., 2010) considerando tres estudios de expresión génica en muestras de tumores de mama y tejido mamario normal: GSE3971 (Zhao et al., 2004), GSE7904 (Richardson et al., 2006) y E-TABM276 (Cheng et al., 2008). De los tres estudios se seleccionaron 5 grupos de muestras: (1) carcinomas ductal de mama vs. tejidos mamarios normales, (2) carcinomas lobular de mama vs. tejidos mamarios normales, (3) carcinomas de mama

vs. tejidos mamarios normales, (4) cáncer de mama basal vs. tejidos mamarios sanos, y (5) cáncer de mama no basal vs. tejidos mamarios sanos. Se determinó las diferencias en expresión génica con significancia estadística entre muestras de tejido tumoral y tejido mamario sano, así como la variación en los niveles de expresión entre ambos tipos de tejidos.

## 5.2.6 Validación en muestras de tejidos y líneas celulares

Se analizó los niveles de metilación de los CpGDM en muestras de tejido mamario sano y tumoral así como líneas celulares, utilizando la base de datos del PEBC (IDIBELL). Específicamente se incluyeron: 67 tumores primarios de cáncer de mama, 18 tejidos sanos de mama y 11 líneas celulares de mama (Tabla Suplementaria 1). En este grupo de tejidos se incluyen tejidos pareados de tumor de mama y tejido mamario sano de 14 pacientes (Tabla Suplementaria 5).

La comparación de los niveles de metilación de los sitios CpG entre los distintos tipos de tejidos se realizó mediante prueba de Wilcoxon rank sum corregido por FDR. Se realizaron análisis de agrupamiento jerárquico de la misma manera que como fue descrito para el estudio de metilación en muestras de ADN de leucocitos (sección 5.2.3).

#### 5.2.7 Secuenciación de ADN tratado con bisulfito

La validación técnica del estatus de metilación determinado por microarreglos fue realizada mediante secuenciación con bisulfito en los laboratorios del PEBC (IDIBELL).

El tratamiento del DNA con bisulfito sódico permite detectar metilación en CpGs basándose en su estado de metilación, ya que el bisulfito sódico deamina los residuos citosina no metilados a uracilos, dejando las 5-metilcitosinas intactas (Clark et al., 1994). En la posterior reacción de PCR, las citosinas serán amplificadas como timinas o citosinas de acuerdo a su estatus de metilación. Esta reacción permite diferenciar ADN metilado del ADN no metilado.

Se amplificaron por PCR los sitios CpG seleccionados como biomarcadores candidatos a partir de ADN de leucocitos tratado con bisulfito de cuatro pacientes con cáncer de mama y cuatro controles. Se utilizó el mismo ADN tratado con bisulfito preparado anteriormente para los ensayos de microarreglos (sección 5.2.2).

Las reacciones de amplificación por PCR, en un volumen final de 15 µl, contenían: buffer de PCR EcoStart 10x (Ecogen), MgCl<sub>2</sub> 50mM (Ecogen), dNTPs 2mM; 1 µM cebadores específicos de la secuencia génica a amplificar y 3 U del enzima ADN polimerasa (EcoStart DNA polimerasa, Ecogen). Los cebadores fueron diseñados para amplificar secuencias de la región de los CpGDM seleccionados y secuencias cercanas al inicio de la transcripción del correspondiente gen. La secuencia de los cebadores empleados para la amplificación de los genes candidatos se recoge en la Tabla 7. Cada reacción de PCR se analizó directamente en geles de agarosa al 2%, teñidos con bromuro de etidio y visualizados bajo iluminación ultravioleta.

abla 7. Cebadores utilizados en la secuenciación del ADN tratado con bisulf	iito.

Sonda	Gen	Región	Cebadores	Secuencia	Producto PCR	
		5′_1 ITP	BS_CYFIP1-F	5' TTGAGAGGAGAATTTGAGAG 3'	263 pb	
Cg20506220 CTFTF1	3-01K	BS_CYFIP1-R	5' ACTAAACACCAAACATAACCTC 3'			
cg1/02/502	ΜΛΡ3ΚΑ	promotor	BS_MAP3K6-F	5' GTTTAGGGTGTAGGTTTTTTT 3'	309 nh	
Cg14024302 MAP3N0		promotor	BS_MAP3K6-R	5' TAAAACTCAACCTCTCCCC 3'	303 pb	
cg01229567	MIR2	promotor	BS_MIB2-F	5' TTGTAGGGAAATTTTTTAGGATT 3'	303 nh	
cg19246761	IVIIDZ	promotor	BS_MIB2-R	5' TATCAACCTAAAAAAAAACCCA 3'	303 pu	

Las secuencias génicas amplificadas por PCR fueron extraídas del gel de agarosa mediante el empleo del Kit QIAquick® Gel Extraction (Qiagen). El ADN extraído fue clonado en bacterias competentes de *Escherichia coli* (NovaBlue Singles<sup>TM</sup>) empleando el vector pGEM-T<sup>®</sup>. La estimación del estado de metilación de cada sitio CpG se realizó mediante secuenciación automática de 10 colonias de cada secuencia de estudio (Applied Biosystems). Se comparó la media de metilación de toda la región evaluada para cada gen entre pacientes y controles mediante prueba de Student.

## **5.3 Resultados**

#### 5.3.1 Perfiles de metilación del ADN de leucocitos

Luego de la validación de las muestras, se utilizaron para los análisis estadísticos los perfiles de metilación del ADN de leucocitos de 22 mujeres con cáncer de mama y 10 mujeres controles no afectadas utilizando la plataforma HumanMethylation450 Beadchip de Illumina, que analiza más de 480.000 sitios CpG a lo largo de todo el genoma dando una medida de metilación cuantitativa (Dedeurwaerder et al. 2011; Sandoval et al. 2011). Al filtrar las sondas con baja calidad y aquellas que contenían SNPs en la secuencia de detección, de los 485.577 sitios CpG originalmente impresos en el microarreglo, mantuvimos 459.999 en los análisis subsecuentes.

Al analizar la distribución de los patrones de metilación a lo largo del genoma en los 32 individuos estudiados observamos que la mayoría de los sitios CpG representados en el microarreglo se encuentran metilados (valor  $\beta > 0,7$ ) (Figura 26).



**Figura 26.** Distribución de frecuencia de los niveles de metilación de los sitios CpG representados en el array 450K HumanMethylation BeadChip en el ADN de leucocitos de los 32 individuos analizados.

Resultado similar se observa cuando se analizan los pacientes con cáncer de mama y los controles de manera independiente (Figura 27). Esto es concordante con la representación de los sitios CpG en el microarreglo utilizado (Tabla Suplementaria 6), donde el 55% de los CpGs mapean fuera de regiones reguladoras génicas, con alta frecuencia de secuencia repetitivas generalmente metiladas (Sandoval et al., 2011).



**Figura 27.** Distribución de frecuencia de los niveles de metilación de los sitios CpG representados en el array 450K HumanMethylation BeadChip en el ADN de leucocitos de: (a) controles no afectados y (b) pacientes con cáncer de mama.

La comparación de los valores medios de metilación de la totalidad de los sitios CpGs analizados entre el grupo de pacientes y grupo control presentó una alta correlación (Spearman,  $r^2=0.997$ ,  $p < 2.2 \times 10^{-16}$ , Figura 28), lo que indica que a nivel global los patrones de metilación en muestras de pacientes y controles son muy similares.



**Figura 28.** Nivel de metilación de los sitios CpGs identificados por el ensayo Infinium 450K HumanMethylation. Se representa la media de los valores beta normalizados para cada CpG en pacientes con cáncer de mama (eje x) y grupo control sano (eje y).

A su vez, el agrupamiento de las muestras utilizando los valores de metilación de 45.000 CpGs seleccionados al azar demuestra que no existen diferencias genómicas entre los dos tipos de muestras que pudieran generar diferencias en CpGs específicos entre pacientes y controles, ya que no son capaces de agrupar las muestras de acuerdo a su status con respecto a la enfermedad (Figura 29).



**Figura 29.** Agrupamiento jerárquico de 45.000 CpGs elegidos al azar en 22 pacientes con cáncer de mama (KP) y 10 controles sanos (KC) analizados en la plataforma Infinium HumanMethylation450 BeadChip. Se utilizó el método de aglomeración "completo" usando distancias "Euclidianas". El nivel de metilación de los CpGs está codificado en colores (verde: menor nivel de metilación; rojo: mayor nivel de metilación).

## 5.3.2 Los análisis de agrupamientos no supervisados revelan un perfil de metilación del ADN específico de pacientes con cáncer de mama.

Al comparar los niveles de metilación de los sitios CpGs entre pacientes con cáncer de mama e individuos controles, detectamos 77 posiciones CpGs diferencialmente metiladas que permiten agrupar a los pacientes y controles de manera separada (Figuras 30, 31 y 32). Estos sitios CpGs diferenciales fueron denominados CpGDM.

El gráfico del análisis de escalamiento multidimensional (MDS) de los valores beta de metilación de los 77 CpGDMs claramente discrimina entra las muestras de pacientes con cáncer de mama y el grupo control en la dimensión 1 (Figura 30).



**Figura 30.** Análisis de escalamiento multidimensional (MDS) de los valores de metilación de los CpGDMs en leucocitos de 22 pacientes con cáncer de mama y 10 controles no afectadas. Las pacientes con cáncer están representadas con color azul, y los controles con color rojo.

Por otro lado, el análisis de agrupamiento jerárquico de las muestras identifica dos grupos principales, separando a las pacientes con cáncer de las mujeres controles, y 3 subgrupos dentro de las muestras de las pacientes con cáncer (Figuras 31 y 32). La fuerza estadística del agrupamiento de las muestras analizando los 77 CpGDMs fue evaluada mediante *bootstrap* de 10.000 remuestreos (Figura 32). Estos resultados indican la existencia de diferencias de metilación CpG específicas entre las muestras de

pacientes con cáncer y controles a nivel de los leucocitos, que pueden ser visualizadas fácilmente mediante estas técnicas de clasificación no supervisada. Dado que en el grupo de pacientes se detecta 3 subagrupamientos definidos (Figuras 31 y 32), analizamos si estos subgrupos estaban asociados a variables epidemiológicas o característica tumoral (Tabla Suplementaria 4). Sin embargo, no encontramos asociación en las muestras de las pacientes con cáncer entre la edad, estatus de fumador, ancestralidad, estadio tumoral, tipo histológico de tumor, receptores hormonales y los subgrupos derivados del análisis de agrupamiento.



**Figura 31.** Agrupamiento jerárquico de los CpGDM en 22 pacientes con cáncer de mama (KP) y 10 controles sanos (KC) analizados en la plataforma Infinium HumanMethylation 450K BeadChip. Se utilizó el método de aglomeración "completo" y distancias euclidianas. Los CpGDMs fueron ordenados por la diferencia de las medias de los valores betas entre pacientes y controles. El nivel de metilación de los CpGDMs está codificado en colores (verde: menor nivel de metilación; rojo: mayor nivel de metilación).

Cluster dendrogram with AU/BP values (%)



**Figura 32.** Dendograma basado en agrupamiento jerárquico mediante boostrap de 10.000 remuestreos, para los valores de metilación de los 77 CpGDM en pacientes con cáncer de mama (KP) y controles (KC). En los puntos de ramificación se indican las probabilidades asociadas AU/BP. La estructura del dendograma es igual al dispuesto en el *heatmap* de la Figura 31.

Con el fin de demostrar que el agrupamiento entre pacientes y controles utilizando los datos de metilación de los 77 CpGDMs no se había obtenido al azar, se realizó un análisis de agrupamiento jerárquico (método de aglomeración "completo" y distancias euclidianas) tomando remuestreos (n=500) de 77 CpGs al azar (del total de 459.999 sitios CpG incluidos inicialmente en los análisis). En ninguno de los *clusters* obtenidos se observaron agrupamientos separados de los pacientes y los controles, respectivamente.

A pesar de ser un número pequeño de muestras, evaluamos el efecto de potenciales confusores en nuestro análisis. Para ello realizamos una regresión lineal para el estatus de la enfermedad (paciente/control) y nivel de metilación de cada CpGDMs ajustada por edad, IMC y estatus de fumadora, y 48 de los CpGDMs continuaron siendo diferenciales entre ambos grupos con significancia estadística (p < 0.05) (Tabla 8).

Todos los CpGDMs y sus características se listan en la Tabla 8. El patrón de metilación de los CpGDMs difiere levemente del presentado por la totalidad de los CpG evaluados por el microarreglo 450K, ya que sólo el 36,4% de los CpGDM están metilados ( $\beta > 0.7$ ), mientras que la mayoría (58,4%) se encuentran hemi-metilados (valores de metilación entre 0,3 y 0,7) en las muestras analizadas (Tabla Suplementaria 7). El 97% (75 de 77) de los CpGDMs presentan menor nivel de metilación en los pacientes con cáncer, con solo el 3% ganando metilación (Figura 33), y 7 CpGDMs presentan una diferencia entre pacientes y controles de la media de los valores beta > 10%.



**Figura 33.** Diferencias en la media de los valores betas entre pacientes con cáncer de mama y controles para los 77 sitios CpG diferencialmente metilados identificados por prueba de Wilcoxon (p<0.05).

**Tabla 8.** Características de los 77 CpG diferencialmente metilados entre pacientes con cáncer de mama y grupo control no afectado. CG: identificación de la sonda, CHR: cromosoma, TSS1500: a 1500 pb del sitio de inicio de la transcripción, TSS200: a 200 pb del sitio de inicio de la transcripción, UTR: región no traducida, BMI: índice de masa corporal.

CG	СПВ		Contexto	Contexto isla CpG	Media pacientes	Media Controles	Delta media	Wilcoxon, FDR	P valor corregido por
00			genomico			0.9107		0.0404	
Cg01015663		ICEAS	DOUY	open sea	0,7431	0,0197	-0,0766	0,0494	0,0216
cg01229567	1	MIB2	1881500	shore	0,4977	0,6160	-0,1184	0,0494	0,9984
cg04400047	1	UBIAD1	TSS1500	shore	0,5763	0,6289	-0,0526	0,0494	0,0293
cg06432479	1	TAL1	body	island	0,3252	0,3909	-0,0657	0,0494	0,0302
cg10159215	1	LPPR5	TSS200	island	0,4070	0,5084	-0,1014	0,0498	0,2749
cg14024502	1	MAP3K6	TSS1500	shore	0,9446	0,9526	-0,0080	0,0498	0,0400
cg15452381	1	HIVEP3	5'UTR	open sea	0,9057	0,9254	-0,0196	0,0498	0,0120
cg16727538	1	C1orf213		shore	0,3201	0,3915	-0,0714	0,0494	0,0188
cg19246761	1	MIB2	TSS1500	shore	0,4267	0,5383	-0,1116	0,0494	0,0482
cg26251270	1	GPX7	TSS1500	shore	0,2750	0,3321	-0,0571	0,0494	0,0135
cg26750487	1	CR1L	TSS200	island	0,3739	0,4587	-0,0848	0,0494	0,0359
cg02537909	2	COBLL1	body	island	0,3697	0,4301	-0,0604	0,0494	0,0321
cg02738156	2		intergenic	open sea	0,4604	0,5389	-0,0785	0,0494	0,0707
cg09408768	2	KCNJ3	TSS200	Island	0,5073	0,5694	-0,0621	0,0494	0,0355
cg17165836	2	AFF3	body	open sea	0,5426	0,6127	-0,0701	0,0494	0,0902
cg24130711	2	CCDC85A	Body	shore	0,3266	0,3807	-0,0541	0,0494	0,0679
cg26175971	2	CYP27A1	Body	Island	0,4551	0,5070	-0,0520	0,0494	0,0390
cg26874367	2		intergenic	open sea	0,3868	0,4773	-0,0904	0,0494	0,0219
cg01615258	3		intergenic	shore	0,3520	0,4041	-0,0521	0,0494	0,0170
cg01814969	3	QARS	Body	shore	0,4735	0,5203	-0,0468	0,0498	0,0342
cg15450445	3		intergenic	open sea	0,9192	0,9333	-0,0141	0,0143	0,9990
cg24840062	3	CDCP1	Body	open sea	0,4984	0,6058	-0,1073	0,0494	0,0453
cg25616514	3		intergenic	open sea	0,7304	0,7618	-0,0315	0,0498	0,0212
cg03002688	4		intergenic	open sea	0,8100	0,8698	-0,0598	0,0498	0,0286
cg18860310	4	SLC10A6	Body	open sea	0,8657	0,9157	-0,0500	0,0494	0,0519
cg26994377	4		intergenic	open sea	0,3417	0,4003	-0,0587	0,0494	0,0727

Tabla	8. (	Contin	uación

		UCSC	Contexto	Contexto	Media	Media	Delta media	Wilcoxon, FDR	P valor corregido por
CG	CHR	REFGENE	genómico	isla CpG	pacientes	Controles	(pacientes-controles)	P valor	edad, BMI y fumar
cg00608540	5	TRIM7	Body	open sea	0,8381	0,8769	-0,0388	0,0498	0,0368
cg01204911	5	ARHGAP26	Body	open sea	0,8387	0,8793	-0,0406	0,0494	0,0202
cg06527989	5	UNC5A	Body	shore	0,3718	0,4453	-0,0735	0,0494	0,0382
cg07475151	5	LOC100268168	Body	open sea	0,5440	0,6446	-0,1005	0,0494	0,0443
cg11953913	5	C5orf32	5'UTR	open sea	0,8184	0,8726	-0,0542	0,0494	0,0984
cg21550107	5		intergenic	open sea	0,8096	0,8357	-0,0261	0,0494	0,2261
cg02174359	6	MRPS18A	Body	shore	0,5117	0,5611	-0,0494	0,0494	0,3319
cg04334016	6	CNKSR3	Body	open sea	0,8170	0,8494	-0,0324	0,0494	0,1056
cg09639771	6		intergenic	open sea	0,7424	0,7867	-0,0442	0,0494	0,0720
cg02377685	7	GBX1	TSS200	Island	0,3191	0,3836	-0,0645	0,0494	0,0534
cg03761471	7	ZYX	TSS1500	shore	0,5038	0,5699	-0,0661	0,0494	0,0262
cg04153882	7	WIPI2	Body		0,8669	0,8010	0,0659	0,0494	0,0362
cg15602580	7	SDK1	Body	shore	0,3081	0,3634	-0,0553	0,0494	0,1723
cg18967180	7	DENND2A	Body		0,4151	0,4831	-0,0680	0,0494	0,1778
cg21252523	7		intergenic	shore	0,8211	0,8605	-0,0394	0,0494	0,0721
cg27403098	7	KIAA1908	Body		0,9375	0,9493	-0,0119	0,0498	0,0435
cg14681767	8	ARHGEF10	5'UTR	Island	0,4175	0,5088	-0,0913	0,0494	0,0812
cg17588094	8		intergenic	shelf	0,9289	0,9456	-0,0167	0,0498	0,0408
cg05616472	9	EHMT1	Body		0,5012	0,5487	-0,0475	0,0494	0,0474
cg14223444	9	TSTD2	5'UTR	shore	0,7831	0,8198	-0,0367	0,0498	0,0587
cg01311537	10	C10orf128	Body	open sea	0,9198	0,9457	-0,0260	0,0494	0,0341
cg08560387	10	TSPAN14	5'UTR		0,3783	0,4925	-0,1142	0,0498	0,0173
cg14770293	10		intergenic	Island	0,5067	0,5830	-0,0763	0,0494	0,0726
cg24168991	10	ITPRIP	5'UTR	shelf	0,8970	0,9277	-0,0308	0,0494	0,0353
cg03611487	11	LOC100126784	Body	shelf	0,8187	0,8672	-0,0486	0,0494	0,0296
cg10141801	11	GUCY2E	TSS200	open sea	0,3802	0,4259	-0,0456	0,0494	0,0533

CG	CHR		Contexto	Contexto isla CpG	Media pacientes	Media Controles	Delta media (pacientes-controles)	Wilcoxon, FDR P valor	P valor corregido por edad_BMI v fumar
00	11		intorgonio		0 7837	0.8380	0.0544	0.0409	0.0244
cg13100902	11	CTI/22		opensea	0,7007	0,0000	-0,0344	0,0498	0,0344
Cg17679104	11	51633	1551500	Island	0,4025	0,4637	-0,0612	0,0494	0,0291
cg22623080	11	AMOTL1	155200	Island	0,3311	0,3748	-0,0437	0,0498	0,0328
cg23460961	11		intergenic	open sea	0,6164	0,6916	-0,0752	0,0494	0,0168
cg04890607	12	HMGA2	Body	open sea	0,5096	0,6427	-0,1331	0,0498	0,0538
cg27292547	12		intergenic	open sea	0,7935	0,8474	-0,0539	0,0498	0,0342
cg11398020	13	KLF5	Body	Island	0,3214	0,3980	-0,0766	0,0494	0,0478
cg01972418	14	PAX9	TSS1500	Island	0,1656	0,2042	-0,0386	0,0498	0,0247
cg18581173	15	CT62	TSS1500	shore	0,4723	0,5396	-0,0673	0,0494	0,9982
cg20172862	15		intergenic	Island	0,0721	0,1030	-0,0309	0,0498	0,9975
cg24359188	15	BUB1B	TSS200		0,5303	0,6061	-0,0758	0,0494	0,0542
cg26568226	15	CYFIP1	5'UTR	Island	0,8514	0,7780	0,0733	0,0494	0,0967
cg04470044	16	WFDC1	Body	shelf	0,8890	0,9149	-0,0259	0,0498	0,0371
cg10155261	16		Body	open sea	0,4262	0,4762	-0,0499	0,0494	0,0316
cg26591162	16	SRL	3'UTR	open sea	0,9212	0,9362	-0,0149	0,0494	0,9975
cg07777703	17	TUBG2	Body	shore	0,5594	0,6085	-0,0491	0,0498	0,1532
cg07848706	17		intergenic	Island	0,3463	0,4184	-0,0722	0,0494	0,0290
cg08960549	17		intergenic	open sea	0,9029	0,9227	-0,0198	0,0494	0,0599
cg09580608	17	GNA13	1stExon	Island	0,3315	0,4167	-0,0852	0,0494	0,0206
cg22163463	17	PITPNM3	Body	shore	0,2386	0,2945	-0,0560	0,0494	0,0186
cg01823541	19	GNG7	5'UTR	Island	0,4138	0,4882	-0,0744	0,0498	0,0381
cg03363633	19	TYROBP	TSS1500	open sea	0,6487	0,7002	-0,0515	0,0494	0,0320
cg22313519	19	KIAA1683	TSS1500		0,7811	0,8196	-0,0386	0,0498	0,0194
cg20477147	20		Body	shore	0,6525	0,7237	-0,0712	0,0494	0,0321
cg26468205	20	PCMTD2	TSS200	Island	0,0547	0,0697	-0,0150	0,0494	0,0270

## Tabla 8. Continuación
#### 5.3.3 Descripción de los CpGDM detectados

De los 77 sitios CpGDM identificados, 59 mapean en 56 genes distintos: 28 en promotores génicos y los restantes en el cuerpo génico (Figura 34a, Tabla 8). De los 18 CpGDMs intergénicos, 3 de ellos se ubican en secuencias repetidas LINEs o LTRs. Considerando la densidad y composición regional de CpG, la mayoría (76.6%) de los CpGDMs están localizados fuera de las islas CpG, con un 24,7% localizados flanqueando las islas (*CpG shores*) (Figura 34b).



**Figura 34.** Distribución genómica de los 77 CpGDMs de acuerdo a su respectiva localización génica (a) y contexto CpG (b).

Analizando separadamente la frecuencia de promotores, cuerpos génicos o regiones intergénicas, no existe un enriquecimiento de los CpGDMs con respecto a su representación en la plataforma del microarreglo (prueba de Fisher, p-valor >0,05). Tampoco existe un enriquecimiento de los CpGDMs ubicados en promotores con respecto a su contexto en isla CpG (Figura 35, Tabla Suplementaria 6).

El análisis de Ontología Génica de los CpGDM reveló un enriquecimiento funcional en procesos biológicos asociados a la regulación del citoesqueleto celular  $(p<1.13 \times 10^{-2})$  y regulación positiva de la transcripción  $(p<3.4 \times 10^{-2})$ .



**Figura 35.** Comparación de los CpGDMs con la distribución genómica y contexto CpG esperado de acuerdo a los CpG impresos en el 450K HumanMethylation BeadChip. TSS1500: 1500 pb del sitio de inicio de la transcripción; TSS200: 200 pb del sitio de inicio de la transcripción, UTR: región no traducida, body: cuerpo del gen.

Diez de los genes detectados con metilación diferencial fueron reportados previamente en G2SBC Database como asociados a susceptibilidad y/o patología de cáncer de mama (*TAL1, MAP3K6, KCNJ3, AFF3, CYP27A1, ZYX, DENND2A, KLF5, CYFIP1* y *TYROBP*), y ampliando la búsqueda de genes previamente asociados a otros tipos de cáncer en Cancer Gene Census, identificamos otros tres genes que se solapan con sitios CpGs identificados como CpGDM (*ARHGAP26, HMGA2* y *BUB1B*) (Tabla 9). Exceptuando *CYFIP1*, todos se encuentran en promedio hipometilados en leucocitos de las pacientes con cáncer de mama en comparación a las mujeres no afectadas.

Gen	Tipo de cáncer	Dirección	Localización	Contexto CpG
TAL1	mama, ALL	hipometilación	cuerpo génico	isla
MAP3K6	mama, ovario	hipometilación	promotor	shore
KCNJ3	mama	hipometilación	promotor	isla
AFF3	mama, ALL, linfoma folicular	hipometilación	cuerpo génico	open sea
CYP27A1	mama, colorectal	hipometilación	cuerpo génico	Isla
ZYX	mama, pulmón, próstata	hipometilación	promotor	Shore
DENND2A	mama	hipometilación	cuerpo génico	open sea
KLF5	mama, esófago, AML, próstata	hipometilación	cuerpo génico	Isla
CYFIP1	mama	hipermetilación	promotor	Isla
TYROBP	mama	hipometilación	promotor	open sea
ARHGAP26	AML, síndrome mieolodisplasia	hipometilación	cuerpo génico	open sea
HMGA2	lipoma, adenoma gl. salival	hipometilación	cuerpo génico	open sea
BUB1B	rhabdomyosarcoma	hipometilación	promotor	open sea

 Tabla 9. CpGDMs en genes previamente asociados a cáncer en Cancer Gene Census y

 G2SBC Database.

Por otro lado, los 7 CpGDMs que presentan una diferencia media de metilación entre pacientes y controles >10% mapean en 5 genes distintos: *MIB2*, *LPPR5*, *CDCP1*, *TSPAN14* y *HMGA2* (Tabla 8).

#### Análisis in sílico de datos de expresión génica disponibles públicamente

Se analizaron las bases de datos públicas de estudios de expresión de ARNm por microarreglos en pacientes con cáncer de mama en comparación con grupo control, tanto a nivel de expresión génica de muestras de sangre (fracción PBMC) como tejido mamario. De manera interesante, en el estudio GSE-27562 donde se presentan datos de expresión génica de PBMC de pacientes con cáncer de mama e individuos control con mamografías normales, 7 de los 56 genes detectados por nosotros con CpGDM presentan diferencias significativas de expresión entre ambos grupos de individuos (Tabla 10).

De igual manera, realizando un meta-análisis en la plataforma NextBio Research considerando tres estudios de expresión génica de muestras de tumores de mama comparados con tejido mamario normal (GSE3971, GSE 7904, E-TABM276), 14 de los genes con CpGDM presentan diferencias significativas entre ambos grupos (Tabla 11).

	Pacientes con tumor de mama benigno vs controles		Pacientes con cáncer de mama vs controles	
	p valor	Incremento de expresión	p valor	Incremento de expresión
GNA13	0.0012	1.69	0.0048	1.47
ARHGAP26	0.0032	1.36	0.0037	1.35
MAP3K6	9 x 10⁻ <sup>6</sup>	-1.22	6.8 x 10 <sup>-6</sup>	-1.23
GBX1	4.4 x 10 <sup>-6</sup>	-1.23	5.1 x 10 <sup>-5</sup>	-1.21
MIB2	2 x 10 <sup>-8</sup>	-1.34	4.1 x10 <sup>-8</sup>	-1.27
TAL1	0.0015	-1.28	0.0059	-1.21
KIAA1683	0.0072	-1.23	0.0047	-1.22

**Tabla 10.** Genes con CpGDMs que presentan diferencias significativas de expresión en PBMC de pacientes con cáncer de mama e individuos control con mamografías normales (GSE27562).

**Tabla 11.** Meta-análisis de expresión génica en muestras de tumores de mama vs tejido mamario sano. Genes con CpGDMs que presentan diferencias significativas de expresión entre muestras de tumores de mama y tejidos mamarios sanos. p valor corregido por FDR, + aumento de expresión génica en tumores, - disminución de la expresión génica en tumores mamarios comparado a tejidos sanos.

Genes	p valor	Variación de expresión
BUB1B	< 3 x 10 <sup>-4</sup>	+
AMOTL1	< 0.001	-
DENND2A	< 0.004	-
GNG7	< 0.01	-
ARHGEF10	$< 9 \times 10^{-4}$	-
TSTD2	< 0.02	-
KIAA1683	$< 2 \times 10^{-4}$	-
KCNJ3	< 0.01	+
MIB2	< 0.007	-
EHMT1	< 0.027	+
MAP3K6	< 3 x 10 <sup>-4</sup>	-
AFF3	< 0.038	-
GNA13	< 0.004	+
KLF5	< 0.030	-

# 5.3.4 Validación del panel de CpGDMs en tejidos de muestreo independiente de origen europeo

Utilizando datos de metilación de CpG obtenidos previamente con la plataforma 450K del banco de tejidos del PEBC correspondientes a población europea, analizamos los niveles de metilación de los 77 CpGDM en tumores primarios de mama y tejidos mamarios sanos. El estudio se completó con la inclusión de líneas celulares comerciales de cáncer de mama.

Al comparar el perfil de metilación de los CpGDMs en las muestras de tejido mamario sano con los observados en leucocitos tanto de pacientes con cáncer de mama como controles, observamos que existe una disminución de los niveles de metilación (valor  $\beta$ ) en el tejido de mama: el 57,1 % de los CpGDMs en tejido mamario sano se encuentran desmetilados ( $\beta < 0,3$ ) en comparación al 5,2% de CpGDMs desmetilados en leucocitos. Situación similar se observa en las muestras de tumor primario de mama, aunque el porcentaje de CpGDMs desmetilados es levemente inferior al tejido sano (Tabla Suplementaria 7). Si bien las muestras de tejidos no corresponden a los mismos individuos que las muestras de sangre, estos resultados sugieren la especificidad de los patrones de metilación en cada tipo celular o tisular.

Sin embargo, a pesar de estas diferencias en los niveles de metilación entre las células sanguíneas y el tejido mamario, los CpGDMs previamente identificadas agrupan la mayoría de las muestras de tejido sano (TMN) diferenciándolas de las muestras tumorales (TP) y líneas celulares (LC) utilizando un análisis de agrupamiento jerárquico (Figura 36).

Esto se visualiza más claramente cuando eliminamos del análisis los datos de las líneas celulares, las cuales tienen características particulares inherentes a su proceso de cultivo y propagación que no son comparables biológicamente desde el punto de vista epigenético con las muestras de tejido mamario tanto sano como tumoral (Figura Suplementaria 1). En la Tabla Suplementaria 8 se describen los valores de metilación promedio de cada uno de los 77 CpGDMs, y la comparación entre tejidos mamarios sanos y tumores primarios.



**Figura 36.** Agrupamiento en dendograma con remuestreo boostrap (n = 10.000) de los 77 CpGDMs entre muestras de tejido mamario sano (TMN), tumores primarios de mama (TP) y líneas celulares de mama (LC). Se utilizó el método de aglomeración "completa". En los puntos de ramificación se indican las probabilidades asociadas AU/BP.

A continuación, analizamos los CpGDM en un grupo de tumores primarios de mama con sus correspondientes tejidos mamarios normales pareados. Encontramos que los 77 CpGDM identificadas eran capaces de separar las muestras de cáncer de las de tejidos sanos utilizando análisis de agrupamiento jerárquico (Figura 37). Únicamente una muestra de tejido sano (N1) fue identificada en el agrupamiento de muestras tumorales, aunque estrechamente cercana a su contrapartida de tejido tumoral. Dado que las muestras de tejido mamario normal fueron extraídas de las inmediaciones cercanas



**Figura 37.** Agrupamiento jerárquico de los 77 CpGDMs en 14 muestras pareadas de tejido mamario normal (N) y tumor primario (T) analizados en la plataforma Infinium HumanMethylation450 BeadChip. Se utilizó el método de aglomeración "completo" usando distancias euclidianas. El nivel de metilación de los CpG está codificado en colores (verde: menor nivel de metilación; rojo: mayor nivel de metilación). El círculo azul indica una muestra de tejido mamario normal dentro del agrupamiento de tejidos tumorales.

al tumor primario, sugiere suponer que esta muestra estuviera contaminada con células tumorales.

Lo que es más interesante, los niveles de metilación de los 77 CpGDMs fueron capaces de separar el tejido tumoral del tejido sano en pacientes con cáncer de mama familiar con mutación en el gen *BRCA2* (pares N5/T5, N6/T6 y N13/T13) (Tabla Suplementaria 5).

De esta manera, validamos el panel de sitios CpG diferencialmente metilados inicialmente detectados en sangre periférica de pacientes con cáncer de mama, en muestras de tejidos de mama provenientes de otra población.

#### 5.3.5 Selección de CpGDM candidatos para su validación clínica

Tomando en cuenta todos los datos analizados en la caracterización de los 77 CpGDM detectados por el análisis de microarreglos, se establecieron los siguientes criterios de selección: *i*) CpGDM ubicados en genes previamente asociados a cáncer, *ii*) CpGDM con mayor diferencias de metilación entre pacientes y controles ( $\Delta\beta \ge 10\%$ ), *iii*) CpGDM hipermetilados en pacientes, *iv*) CpGDM diferencialmente metilados entre tejido mamario tumoral y tejido mamario sano, y con mayor diferencia de metilación entre ambos grupos, *v*) CpGDM ubicados en regiones reguladoras del gen.

De esta manera, se seleccionaron un grupo de CpGDM que cumplían con varios de los criterios de selección descritos más arriba, siendo los primeros 4 CpGDMs seleccionados para su validación: cg26568226 en región promotora de *CYFIP1*, cg14024502 en promotor del gen *MAP3K6* y, cg01229567 y cg19246761 ambas en el promotor de *MIB2*. En la Tabla Suplementaria 9 se resume los criterios que cumplen cada uno de los CpGDMs seleccionados.

Para hacer la validación técnica del estatus de metilación determinado por microarreglos, analizamos en 4 pacientes y 4 controles la metilación de los CpGDMs seleccionados y sus regiones flanqueantes mediante secuenciación del ADN tratado con bisulfito. Dado que las diferencias entre pacientes y controles observadas en el ensayo de microarreglos son muy pequeñas, es difícil visualizar estas mismas diferencias analizando un número limitado de clones por secuenciación de bisulfito. A pesar de ello, cuando comparamos la metilación en los sitios CpG en la región de los CpGDMs seleccionados, se observan diferencias significativas entre pacientes y controles. Evaluando una secuencia de 263 pb donde se incluye el sitio cg26568226 en la región promotora de *CYFIP1*, se observó una clara hipermetilación de esta secuencia en pacientes con cáncer de mama ( $p = 2.22 \times 10^{-12}$ ) (Figura 38). Los resultados para los sitios cg14024502 en el promotor del gen *MAP3K6* y, cg01229567 y cg19246761 ambos en el promotor de *MIB2* se muestran en las Figuras Suplementarias 2 y 3, donde se corrobora una leve hipometilación en ambos casos en los pacientes con cáncer de mama.



**Figura 38.** Secuenciación con bisulfito de los clones individuales a las regiones flanqueantes al sitio CpG evaluado por la sonda cg2658226 en el gen *CYFIP1*. Cada cuadrado corresponde a un sitio CpG en el fragmento amplificado. KC: controles; KP: pacientes. Cuadrados oscuros: metilados; cuadrados blancos: no metilados. Se indica el porcentaje medio de metilación de la región evaluada para cada individuo.

CYFIP1

## 5.4 Discusión

Los pacientes con cáncer de mama localizado tienen una tasa de sobrevida a 5 años de 98%. Sin embargo, cuando son diagnosticados luego de que el tumor sufrió metástasis, la tasa de sobrevida disminuye drásticamente a 27%. Estos resultados marcan el beneficio del tamizaje y la detección temprana. Dado que la sensibilidad de la mamografía es baja (50% en algunos grupos de mujeres), se ha vuelto vital buscar nuevos marcadores para complementar los resultados de la mamografía (Brooks et al., 2009). La dificultad en encontrar marcadores genéticos de riesgo a cáncer de mama se ha ejemplificado recientemente por una serie de grandes estudios de caso-control que han identificado aumentos de poca importancia en el riesgo de cáncer de mama (Ahsan et al., 2014; Campa et al., 2014; Easton et al., 2007; Hicks et al., 2013; Stephens et al., 2012).

Para cáncer de mama esporádico, fueron detectados en tumores primarios una variedad de cambios en la metilación del ADN; sin embargo, los marcadores epigenéticos en fluidos biológicos previamente reportados pierden la sensibilidad y especificidad observada en otros tipos de cáncer (Brennan et al., 2012; Ito et al., 2008; Korshunova et al., 2008; Wojdacz et al., 2011). Esto podría ser debido a las aproximaciones de un solo gen y las tecnologías de baja resolución utilizadas hasta la fecha, que proporcionan visiones limitadas del genoma. Así, recientemente los metilomas de resolución de bases y las plataformas de microarreglos de alta resolución han incrementado nuestro conocimiento sobre del desarrollo embrionario (Doi et al., 2009; Lister et al., 2011) y las enfermedades (Fernandez et al., 2011; Irizarry et al., 2009; Javierre et al., 2010; Lange et al., 2012), incluyendo el cáncer de mama (Heyn et al., 2013b; Hon et al., 2012).

Para intentar sumar información a los esfuerzos anteriores y con el objetivo de buscar diferencias epigenéticas en el ADN de leucocitos para identificar posibles marcadores de riesgo a cáncer de mama en la población uruguaya, utilizamos en este estudio la plataforma de alta resolución Infinium HumanMethylation450 BeadChip. Detectamos 77 sitios CpGs diferencialmente metilados en sangre de pacientes con cáncer de mama en comparación a individuos controles. Este panel de 77 CpGs fue validado en un muestreo independiente de tejidos mamarios pertenecientes a pacientes con cáncer de mama de origen europeo.

#### 5.4.1 Perfil de metilación diferencial en pacientes con cáncer de mama

En el panel de sitios CpGs diferencialmente metilados en pacientes con cáncer de mama, detectamos casi exclusivamente CpGDMs hipometilados en pacientes (75 de 77). Estos resultados están en concordancia con la hipometilación genómica global detectada por HPLC en las mismas muestras de estos pacientes (Capítulo 3). De la misma manera, Heyn y colaboradores (2013b) detectaron CpG diferencialmente metilados en sangre de pares de mellizos discordantes para cáncer de mama, los cuales en su mayoría se encontraban hipometilados en los mellizos afectados por cáncer. En el mismo sentido, se reportó una pérdida global de metilación del ADN en muestras de tumores primarios de mama analizados a nivel de sitios CpGs (Hon et al., 2012). Estos resultados apoyan la sensibilidad de nuestro abordaje basado en el análisis de sitios CpG diferencialmente metilados en sangre.

A pesar que la mayoría de los CpGDM presentan diferencias promedio entre pacientes y controles menores al 10%, estas diferencias son consistentes. La identificación de biomarcadores en sangre tiene el desafío de que no podemos excluir eventos específicos de las células de la sangre, y que los niveles de metilación del ADN tumoral presente en la sangre pueden modificar levemente los perfiles de metilación específicos del ADN de células sanguíneas. Por lo tanto, se espera que las alteraciones presenten cambios de magnitud pequeña, pero consistentes entre pacientes con cáncer y controles. Incluso en el estudio de Heyn y colaboradores (2013b), donde remueven el ruido genético y reducen otras fuentes de confusores al analizar pares de mellizos idénticos discordantes para cáncer de mama, detectan 403 sitios diferencialmente metilados en el ADN de sangre entre mellizos discordantes, todos con diferencias menores al 8% entre ambos grupos. Así, aunque de magnitud pequeña, la integración de múltiples biomarcadores epigenéticos en firmas predictivas pueden ser de alto valor translacional.

Es importante destacar que dentro del panel de CpGDMs detectados en leucocitos, se describen sitios CpG ubicados en genes identificados previamente asociados a cáncer de mama, lo que refuerza la utilidad de este abordaje en la búsqueda de biomarcadores asociados a cáncer de mama en sangre periférica. Incluso, los genes donde mapean 59 CpGDM están enriquecidos en procesos biológicos generalmente alterados durante el desarrollo tumoral como la regulación de la transcripción y regulación del citoesqueleto celular.

El panel de 77 CpGDM detectado, no solo agrupa las muestras de ADN de leucocitos en pacientes y controles, sino que los pacientes quedan subagrupados en 3 clusters bien definidos. Varios estudios de microarreglos demostraron que los tumores de mama pueden ser divididos en al menos 5 subtipos moleculares basados en sus perfiles de expresión génica (Hu et al., 2006; Sørlie et al., 2001). Además, los subtipos moleculares tipo basal, luminal-A y luminal-B presentan perfiles de metilación del ADN específicos (Holm et al., 2010). Con estos datos previos, evaluamos la posible asociación entre los distintos subtipos tumorales, grados y estadíos con los perfiles de metilación de los 77 CpG que separaban a los pacientes en 3 subgrupos, esperando ver reflejado lo que sucede en el tejido tumoral en las muestras de sangre. Sin embargo, no detectamos ninguna asociación significativa de estas variables clínicas que pudieran explicar el agrupamiento, incluso tampoco otras variables epidemiológicas como IMC, edad y estatus de fumador.

Por otro lado, el panel de CpGDMs fue validado en un muestreo de tejidos de tumores primarios de mama y sus correspondientes tejidos sanos pareados. Aun detectando niveles más bajo de metilación de los CpGDMs tanto en el tejido mamario sano como tumoral en comparación a los niveles de los mismos sitios CpG en sangre, el panel de 77 CpGDMs clasifica los tejidos sanos y sus contrapartidas tumorales en grupos separados.

Cada tejido posee una firma epigenética única, que muchas veces refleja su funciones diferenciales (Armstrong et al., 2014; Fernandez et al., 2011). A pesar de que la mayor variación en la metilación del ADN se observa entre tejidos, las diferencias interindividuales en la metilación del ADN en tejidos internos se correlacionan con las de las células sanguíneas para un grupo de sitios CpG (Slieker et al., 2013). Esto refuerza la teoría de la existencia de variación en la metilación del ADN a nivel sistémico que pudiera estar asociada a predisposición a ciertas enfermedades. Debido a que en nuestro estudio las muestras de ADN de leucocitos no corresponden a las mismas mujeres de quienes proceden las muestras de tejido mamario, no podemos

afirmar que exista una correlación en los niveles de metilación de los CpGDM entre ambos tejidos. Sin embargo, el panel de 77 CpGDMs detectados en leucocitos de pacientes con cáncer fue validado en tejidos mamarios de mujeres de origen europeo, diferenciando el tejido sano del tumoral en cada caso. Esto sugiere el potencial de este panel como marcador de cáncer de mama.

Aún más importante, este panel de CpGDMs clasifica a los tejidos sanos y tumorales incluso en pacientes con cáncer de mama familiar, portadoras de mutación en el gen BRCA2. Actualmente es bien conocido que alteraciones epigenéticas ocurren frecuentemente en el cáncer de mama esporádico, pero poco se conoce sobre las alteraciones epigenéticas asociadas a tumores de mama familiares. Un estudio previo sugirió que los cánceres de mama hereditarios poseían metilación similar a los tumores esporádico en 10 genes candidatos (Esteller et al., 2001). Sin embargo, un estudio más reciente sugirió que los tumores BRCA1 presentan menor nivel de metilación que los tumores de mama esporádicos en 11 genes candidatos (cinco superpuestos con el estudio anterior) (Suijkerbuijk et al., 2008). Luego, Flanagan y colaboradores (2010) demostraron, mediante un estudio genómico, que los perfiles de metilación de los cánceres de mama familiares estaban definidos por su estatus de metilación, con niveles menores de metilación en tumores BRCA1. Esto es concordante con el papel que juega BRCA1, conocido hoy en día, en la metilación global del ADN regulando la expresión de DNMT1(Shukla et al., 2010). Sin embargo, aún es desconocido si los perfiles de metilación de tumores BRCA2 son iguales o diferentes al de los tumores de mama esporádicos. De esta manera, nuestros resultados sugieren la existencia de alteraciones epigenéticas en los tumores BRCA2 ya que permiten distinguirlos de tejido normal mediante la detección de variabilidad a nivel de la metilación del ADN, y plantea la posible potencialidad del uso de nuestro panel de CpGDM tanto para diferenciar casos de cáncer de mama esporádico como tumores portadores de mutación en BRCA2.

Cabe destacar que las muestras de ADN de leucocitos de pacientes y controles estaban pareadas por ancestralidad, lo que se espera que no haya influencia de la ancestría genética en los perfiles de metilación de los individuos y por lo tanto en la detección de sitios CpG diferencialmente metilados en pacientes con cáncer. Al eliminar el componente ancestral como confusor, facilita la validación del panel de CpGDM en muestras de otras poblaciones, como las utilizadas para la validación en tejidos.

147

#### 5.4.2 Sitios CpGDM seleccionados para su validación técnica y clínica

Moviéndonos desde una estrategia de identificación genómica a una fase de validación gen específica, seleccionamos 4 sitios CpGDMs para analizar en más detalle, que mapean en 3 genes: *CYFIP1*, *MAP3K6* y *MIB2*. La selección se realizó utilizando criterios basados en la caracterización previa realizada a todo el panel de CpGDMs en cuanto a las diferencias en los perfiles de metilación entre pacientes y controles, contexto genético y contexto isla CpG, y análisis *in situ* de datos obtenidos utilizando bases de datos públicas en cuanto a su asociación a cáncer, niveles de expresión en tejidos mamarios sanos y tumorales así como en células sanguíneas. La metilación diferencial en la región de los 4 CpGDM fue validada técnicamente por secuenciación con bisulfito de sodio.

Específicamente el sitio CpGDM en el gen *CYFIP1* se ubica en una isla CpG en la región 5' del gen, y es uno de los dos CpGDMs hipermetilados en leucocitos de pacientes con cáncer de mama. Este gen, que codifica una proteína citoplasmática de unión a FMRP y cuya función está asociada a la regulación del citoesqueleto y regulación de la traducción, se ubica en la región 15q11.2 adyacente a la región improntada asociada al síndrome Prader-Willi/Angelman (Chai et al., 2003). Esta región cromosómica está ligada a varios desórdenes neurológicos, incluyendo retardo mental, autismo y esquizofrenia (De Rubeis et al., 2013). Por otro lado, este gen se encuentra frecuentemente deleccionado en cánceres epiteliales, y se ha reportado la reducción de su expresión durante la invasión tumoral, implicando un mal pronóstico en tumores epiteliales (Silva et al., 2009). Mediante análisis *in situ*, no detectamos diferencias en los niveles de expresión de *CYFIP1* al comparar tejidos mamarios sanos con tumores primarios, lo que sugiere un papel de este gen en etapas más avanzadas del desarrollo tumoral.

Con respecto al segundo sitio CpGDMs seleccionado para su validación, este mapea en el promotor del gen *MAP3K6*, el cual codifica una quinasa involucrada en la activación de MAP quinasas. El sitio CpG diferencialmente metilado en leucocitos detectado en el promotor del gen MAP3K6 en nuestro trabajo, se encuentra metilado tanto en pacientes como en controles ( $\beta > 0.7$ ), aunque con un nivel de metilación menor y significativo en pacientes con cáncer de mama. MAP3K6 fue originalmente identificada como un miembro de la familia de serin/treonin quinasas por su interacción con MAP3K5, una proteína que también activa c-Jun quinasa y p38 (Wang et al., 1998). Años después se demostró que MAP3K6 participa en la regulación del factor de crecimiento endotelial vascular (VEGF), el cual es un regulador crítico en la vascularización tumoral. Las células tumorales con baja expresión de *MAP3K6* muestran una supresión significativa en el crecimiento tumoral *in vivo*, el cual está acompañado de una represión de la formación de vasos sanguíneos en los tumores (Eto et al., 2009).

Los últimos dos sitios CpGDMs seleccionados para iniciar la validación del panel mapean ambos en el promotor del gen *MIB2* (también conocido como skeletrophin). *MIB2* codifica una ubiquitin ligasa y está localizado en una región cuya pérdida de heterocigocidad está asociada a la tumorigenicidad e invasión en melanoma maligno. Específicamente, melanoma maligno invasivo no expresa esta proteína, mientras la mayoría de los nevos benignos si lo hacen (Takeuchi et al., 2006). De manera opuesta, el gen *MIB2* se encuentra sobreexpresado en células plasmáticas malignas en mieloma múltiple y se especula que dicha sobreexpresión está dirigida por hipometilación de su promotor (Takeuchi et al., 2005). En este contexto, los dos sitios CpGDMs detectados en sangre en nuestro estudio presentaron niveles medios de metilación (hemimetilados) tanto en pacientes como controles. Sin embargo en las muestras de los pacientes con cáncer de mama se detectó un hipometilación significativa de más del 10% en comparación a las muestras de los controles. Igualmente, el rol que juega el gen *MIB2* en los distintos tipos de cáncer tiene que ser elucidado ya que no es claro su papel en el desarrollo tumoral.

#### 5.4.3 Conclusiones

Estudiando ADN de leucocitos de 22 pacientes con cáncer de mama y 10 controles no afectados, el análisis de metilación del ADN de alta resolución determinó un grupo de sitios CpGs diferencialmente metilados que incluía sitos en genes asociados a cáncer conocidos y nuevos candidatos con posibles implicancias en el riesgo al desarrollo tumoral. Este panel de sitios CpGs fue validado en muestras de tejidos tumorales de una cohorte independiente de origen europeo. Actualmente en nuestro laboratorio estamos comenzando la validación de los CpGDMs seleccionados como candidatos en un muestreo mayor e independiente mediante MS-HRM.

Por primera vez en nuestro país, se describe la búsqueda y detección de un panel de posibles marcadores de metilación del ADN en leucocitos asociados a riesgo a cáncer de mama en nuestra población.



# **6.1 Conclusiones finales**

Al inicio del trabajo de tesis nos planteamos la pregunta de si era posible detectar metilación diferencial en el ADN de leucocitos que se pudiese asociar con la presencia de cáncer. La hipótesis de partida se basa en la existencia de cambios epigenéticos a nivel sistémico que pudieran predisponer al desarrollo tumoral. Teniendo esta pregunta en mente, nos planteamos un diseño experimental utilizando dos acercamientos metodológicos complementarios. En primer lugar, cuantificamos el nivel global de citosinas metiladas por Cromatografía líquida de alta resolución (HPLC-ESI/MS). Confirmada la existencia de diferencias en los niveles genómicos de metilación entre pacientes y controles, realizamos estudios para localizar las regiones genómicas con metilación diferencial sitio-específica empleando un análisis de alta resolución a nivel de sitios CpG basado en tecnología de microarreglos. De esta manera aplicamos un abanico amplio de metodologías, incluyendo tecnologías innovadoras para este tipo de estudios como los ensayos de microarreglos de metilación de alta resolución.

De los resultados descritos en esta tesis se desprenden las siguientes conclusiones finales:

1) Detectamos la presencia de una hipometilación genómica global en sangre periférica de pacientes con cáncer de mama y melanoma cutáneo esporádico.

2) El contexto genético de los individuos tiene un efecto sobre la variabilidad epigenética a nivel global en nuestra población, dada por la ancestría genética de

los individuos y/o variantes genéticas en genes asociados directa o indirectamente con los mecanismos epigenéticos.

**3**) Detectamos una asociación inversa entre el componente genético ancestral africano y el porcentaje de metilación global del ADN de leucocitos en pacientes con cáncer de mama.

4) Detectamos una asociación positiva entre el genotipo CC del polimorfismo de nucleótido simple del gen *MUS81* y los niveles de metilación global en los individuos controles no afectados por cáncer de mama, sugiriendo un papel protector contra la inestabilidad genómica.

**5**) Identificamos por primera vez metilación diferencial sitio-específica en el ADN genómico de leucocitos de pacientes con melanoma cutáneo maligno, confirmando el potencial de la técnica de AIMS para la detección de biomarcadores.

**6**) Validamos la metilación diferencial asociada al cáncer de melanoma en dos regiones genómicas correspondientes a los genes neurotrimina (*NTM*) y la proteína transmembrana 184A (*TMEM184A*) en tejidos de una población independiente compuesta por individuos europeos, proponiendo el estudio de la metilación de *NTM* y *TMEM184A* para validar su potencial valor como biomarcador de riesgo de desarrollo de cáncer de melanoma.

**7**) Basado en el análisis de más de 485.000 sitios de metilación específica a lo largo de todo el genoma, proponemos un panel de 77 sitios CpG diferencialmente metilados en una muestra de pacientes con cáncer de mama esporádico para estudiar su valor potencial como biomarcadores de riesgo de desarrollo de cáncer de mama.

8) Desde tres aproximaciones diferentes identificamos metilación diferencial en el ADN de leucocitos de pacientes con cáncer de mama o melanoma esporádicos, que permiten pensar en la existencia de variabilidad en la metilación del ADN sistémica que está relacionada con la susceptibilidad al desarrollo tumoral.

## 6.2 Perspectivas del trabajo

Este trabajo es solo el inicio en la búsqueda de marcadores epigenéticos de riesgo al cáncer de mama y melanoma en la población uruguaya. Aunque los resultados son promisorios, será necesaria más investigación para determinar las causas de las diferencias de metilación detectadas y si la metilación diferencial en el ADN de leucocitos de pacientes con cáncer de mama y melanoma esporádico puede ser validado como herramienta de tamizaje o diagnóstica en nuestra población.

También sería importante evaluar si pacientes con lesiones precancerosas o benignas exhiben patrones de metilación similares a los descritos tanto a nivel global como sitio-específico en pacientes con cáncer de mama y melanoma.

Debido a que la metilación del ADN se plantea como un predictor para el desarrollo de cáncer, la variabilidad epigenética observada en leucocitos implica que las diferencias en la metilación del ADN tanto a nivel global como secuencia específica colectivamente reflejan la condición sistémica que puede favorecer el desarrollo tumoral. En este sentido, sería de gran interés realizar un seguimiento a los individuos por largos períodos de tiempo y ampliar el muestreo poblacional para poder evaluar la variabilidad en la metilación del ADN de acuerdo a diferentes tipos de cáncer así como la capacidad de detectar la población de riesgo antes de la manifestación clínica de la enfermedad.

En resumen, en este trabajo se describe por primera vez la búsqueda de biomarcadores epigenéticos candidatos en leucocitos de sangre periférica asociados a tumores sólidos en la población uruguaya. Confirmamos diferencias en la metilación del ADN entre pacientes con cáncer de mama y melanoma cutáneo e individuos no afectados, y pudimos identificar un grupo de sitios CpG diferencialmente metilados que parecen ser altamente indicativos de la presencia de cáncer. La validación de estos resultados en un muestreo mayor e independiente en un futuro próximo, abrirá las puertas a la evaluación de paneles de biomarcadores propios de nuestra población.

# Bibliografía

Adkins, R.M., Krushkal, J., Tylavsky, F.A., and Thomas, F. (2011). Racial differences in gene-specific DNA methylation levels are present at birth. *Birth Defects Res.* 736, 728–736.

Ahsan, H., Halpern, J., Kibriya, M.G., Pierce, B.L., Tong, L., Gamazon, E.R., McGuire, V., Felberg, A., Shi, J., Jasmine, F., et al. (2014). A Genome-wide association study of earlyonset breast cancer identifies PFKM as a novel breast cancer gene and supports a common genetic spectrum for breast cancer at any age. *Cancer Epidemiol. Biomarkers Prev.* DOI: 10.1158/1055-9965.

Ally, M.S., Al-Ghnaniem, R., and Pufulete, M. (2009). The relationship between genespecific DNA methylation in leukocytes and normal colorectal mucosa in subjects with and without colorectal tumors. *Cancer Epidemiol. Biomarkers Prev.* 18, 922–928.

Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Bonnen, P.E., de Bakker, P.I.W., Deloukas, P., Gabriel, S.B., et al. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature* 467, 52–58.

Amos, C.I., Wang, L.-E., Lee, J.E., Gershenwald, J.E., Chen, W. V, Fang, S., Kosoy, R., Zhang, M., Qureshi, A. a, Vattathil, S., et al. (2011). Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum. Mol. Genet.* 20, 5012–5023.

Armstrong, D.A., Lesseur, C., Conradt, E., Lester, B.M., and Marsit, C.J. (2014). Global and gene-specific DNA methylation across multiple tissues in early infancy: implications for children's health research. *FASEB J.* DOI: 10.1096/fj.13-238402.

Ballestar, E., and Esteller, M. (2008). Epigenetic gene regulation in cancer. *Adv. Genet.* 61, 247–267.

Barnes, S.K., and Ozanne, S.E. (2011). Pathways linking the early environment to long-term health and lifespan. *Prog. Biophys. Mol. Biol.* 106, 323–336.

Barrett, J.H., Iles, M.M., Harland, M., Taylor, J.C., Aitken, J.F., Andresen, A., Akslen, L.A., Armstrong, B.K., Avril, M., Bakker, B., et al. (2012). Genome-wide association study identifies three new melanoma susceptibility loci. *Nat. Genet.* 43, 1108–1113.

Barrios, E., Vassallo, J., Alonso, R., Garay, M., and Musetti, C. (2010). III Atlas de incidencia del cáncer en Uruguay 2002-2006 (Montevideo).

Baylin, S., and Bestor, T.H. (2002). Altered methylation patterns in cancer cell genomes: cause or consequence? *Cancer Cell* 1, 299–305.

Bell, J.T., Tsai, P., Yang, T., Pidsley, R., Nisbet, J., Glass, D., Mangino, M., Zhai, G., Zhang, F., Valdes, A., et al. (2012). Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet.* 8, e1002629.

Berdasco, M., Fraga, M.F., and Esteller, M. (2009). Quantification of global DNA methylation by capillary electrophoresis and mass spectrometry. *Methods Mol. Biol.* 507, 23–34.

Berdasco, M., and Esteller, M. (2010). Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev. Cell* 19, 698–711.

Berdasco, M., and Esteller, M. (2012). Hot topics in epigenetic mechanisms of aging: 2011. *Aging Cell* 11, 181–186.

Berdasco, M., and Esteller, M. (2013). Genetic syndromes caused by mutations in epigenetic genes. *Hum. Genet.* 132, 359–383.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.

Best, D., and Adams, I.R. (2009). Sdmg1 is a component of secretory granules in mouse secretory exocrine tissues. *Dev. Dyn.* 238, 223–231.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. 16, 6–21.

Bird, A. (2007). Perceptions of epigenetics. Nature 447, 396–398.

Bjornsson, H.T., Fallin, M.D., and Feinberg, A.P. (2004). An integrated epigenetic and genetic approach to common human disease. *Trends Genet*. 20, 350–358.

Bjornsson, H.T., Sigurdsson, M.I., Fallin, M.D., Irizarry, R.A., Aspelund, T., Cui, H., Yu, W., Rongione, M.A., Ekström, T.J., Harris, T.B., et al. (2008). Intra-individual change over time in DNA methylation with familial clustering. *JAMA* 299, 2877–2883.

Bocker, M.T., Hellwig, I., Breiling, A., Eckstein, V., Ho, A.D., and Lyko, F. (2011). Genome-wide promoter DNA methylation dynamics of human hematopoietic progenitor cells during differentiation and aging. Blood 117, e182–9.

Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular signals of epigenetic states. *Science* 330, 612–616.

Bonilla, C., Bertoni, B., Sans, M., Hidalgo, P.C., Artagaveytia, N., Figueiro, G., Cappetta, M., Achermann, E., Barreto, I., Egaña, A., et al. (2010). Ancestría genética y cáncer de mama esporádico en una muestra de mujeres uruguayas. In 110 Congreso Uruguayo de Oncología, (Montevideo).

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science* 294, 2536–2539.

Brazma, A., Parkinson, H., Sarkans, U., Shojatalab, M., Vilo, J., Abeygunawardena, N., Holloway, E., Kapushesky, M., Kemmeren, P., Lara, G.G., et al. (2003). ArrayExpress--a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res.* 31, 68–71.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1984). Classification and regression trees. Chapman and Hall/CRC; 1st edition.

Brennan, K., and Flanagan, J.M. (2012). Is there a link between genome-wide hypomethylation in blood and cancer risk? *Cancer Prev. Res.* 5, 1345–1357.

Brennan, K., Garcia-Closas, M., Orr, N., Fletcher, O., Jones, M., Ashworth, A., Swerdlow, A., Thorne, H., Riboli, E., Vineis, P., et al. (2012). Intragenic ATM methylation in peripheral blood DNA as a biomarker of breast cancer risk. *Cancer Res.* 72, 2304–2313.

Brooks, J., Cairns, P., and Zeleniuch-jacquotte, A. (2009). Promoter methylation and the detection of breast cancer. *Cancer Causes Control* 20, 1539–1550.

Campa, D., Barrdahl, M., Tsilidis, K.K., Severi, G., Diver, W.R., Siddiq, A., Chanock, S., Hoover, R.N., Ziegler, R.G., Berg, C.D., et al. (2014). A genome-wide "pleiotropy scan" does not identify new susceptibility Loci for estrogen receptor negative breast cancer. *PLoS One* 9, e85955.

Cash, H.L., Tao, L., Yuan, J.M., Marsit, C.J., Houseman, E.A., Xiang, Y.B., Gao, Y.T., Nelson, H.H., and Kelsey, K.T. (2011). LINE-1 hypomethylation is associated with bladder cancer risk among nonsmoking Chinese. *Int. J. Cancer* 130, 1151–1159.

Chai, J.-H., Locke, D.P., Greally, J.M., Knoll, J.H.M., Ohta, T., Dunai, J., Yavor, A., Eichler, E.E., and Nicholls, R.D. (2003). Identification of four highly conserved genes between breakpoint hotspots BP1 and BP2 of the Prader-Willi/Angelman syndromes deletion region that have undergone evolutionary transposition mediated by flanking duplicons. *Am. J. Hum. Genet.* 73, 898–925.

Chan, T.L., Yuen, S.T., Kong, C.K., Chan, Y.W., Chan, A.S.Y., Ng, W.F., Tsui, W.Y., Lo, M.W.S., Tam, W.Y., Li, V.S.W., et al. (2006). Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nat. Genet.* 38, 1178–1183.

Chen, J., Lui, W.O., Vos, M.D., Clark, G.J., Takahashi, M., Schoumans, J., Khoo, S.K., Petillo, D., Lavery, T., and Sugimura, J. (2003). The t(1;3) breakpoint-spanning genes LSAMP and NORE1 are involved in clear cell renal cell carcinomas. *Cancer Cell* 4, 405–413.

Cheng, A.S.L., Culhane, A.C., Chan, M.W.Y., Venkataramu, C.R., Ehrich, M., Nasir, A., Rodriguez, B.A.T., Liu, J., Yan, P.S., Quackenbush, J., et al. (2008). Epithelial progeny of estrogen-exposed breast progenitor cells display a cancer-like methylome. *Cancer Res.* 68, 1786–1796.

Choi, J.D., and Lee, J.-S. (2013). Interplay between Epigenetics and genetics in cancer. *Genomics Inform*. 11, 164–173.

Choi, J.-Y., James, S.R., Link, P. a, McCann, S.E., Hong, C.-C., Davis, W., Nesline, M.K., Ambrosone, C.B., and Karpf, A.R. (2009). Association between global DNA hypomethylation in leukocytes and risk of breast cancer. *Carcinogenesis* 30, 1889–1897.

Christensen, B.C., Houseman, E.A., Marsit, C.J., Zheng, S., Wrensch, M.R., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Padbury, J.F., Bueno, R., et al. (2009). Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genet.* 5, e1000602.

Clark, S.J., Harrison, J., Paul, C.L., and Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* 22, 2990–2997.

Conway, K., Edmiston, S.N., Khondker, Z.S., Groben, P. a, Zhou, X., Chu, H., Kuan, P.F., Hao, H., Carson, C., Berwick, M., et al. (2011). DNA-methylation profiling distinguishes malignant melanomas from benign nevi. *Pigment Cell Melanoma Res.* 24, 352–360.

Cui, Y., Ying, Y., van Hasselt, A., Ng, K.M., Yu, J., Zhang, Q., Jin, J., Liu, D., Rhim, J.S., Rha, S.Y., et al. (2008). OPCML is a broad tumor suppressor for multiple carcinomas and lymphomas with frequently epigenetic inactivation. *PLoS One* 3, e2990.

Czyz, W., Morahan, J.M., Ebers, G.C., and Ramagopalan, S.V (2012). Genetic, environmental and stochastic factors in monozygotic twin discordance with a focus on epigenetic differences. *BMC Med.* 10, 93–115.

Dahl, C., and Guldberg, P. (2007). The genome and epigenome of malignant melanoma. *APMIS* 115, 1161–1176.

Dahl, C., Grønbæk, K., and Guldberg, P. (2011). Advances in DNA methylation: 5hydroxymethylcytosine revisited. *Clin. Chim. Acta* 412, 831–836.

Daniel, F.I., Cherubini, K., Yurgel, L.S., de Figueiredo, M.A.Z., and Salum, F.G. (2011). The role of epigenetic transcription repression and DNA methyltransferases in cancer. *Cancer* 117, 677–687.

Da Silveira, L. (2012). Puesta a punto y evaluación de la técnica de amplificación de sitios intermetilados. Tesis de Licenciatura, Facultad de Ciencias - Universidad de la República.

Das, P.M., Ramachandran, K., Vanwert, J., Ferdinand, L., Gopisetty, G., Reis, I.M., and Singal, R. (2006). Methylation mediated silencing of TMS1/ASC gene in prostate cancer. *Mol. Cancer* 5, 28.

De Rubeis, S., Pasciuto, E., Li, K.W., Fernández, E., Di Marino, D., Buzzi, A., Ostroff, L.E., Klann, E., Zwartkruis, F.J.T., Komiyama, N.H., et al. (2013). CYFIP1 coordinates mRNA translation and cytoskeleton remodeling to ensure proper dendritic spine formation. *Neuron* 79, 1169–1182.

Dedeurwaerder, S., Defrance, M., Calonne, E., Denis, H., Sotiriou, C., and Fuks, F. (2011). Evaluation of the Infinium Methylation 450K technology. *Epigenomics* 3, 771–784.

Delgado, L., Fernández, G., González, A., Bressac-de Paillerets, B., Gualco, G., Bombled, J., Cataldi, S., Sabini, G., Roca, R., and Musé, I.M. (2002). Hereditary breast cancer associated with a germline BRCA2 mutation in identical female twins with similar disease expression. *Cancer Genet. Cytogenet.* 133, 24–28.

Delgado, L., Fernández, G., Grotiuz, G., Cataldi, S., González, A., Lluveras, N., Heguaburu, M., Fresco, R., Lens, D., Sabini, G., et al. (2011). BRCA1 and BRCA2 germline mutations in Uruguayan breast and breast-ovarian cancer families. Identification of novel mutations and unclassified variants. *Breast Cancer Res. Treat.* 128, 211–218.

Delgado-Cruzata, L., Wu, H.-C., Perrin, M., Liao, Y., Kappil, M.A., Ferris, J.S., Flom, J.D., Yazici, H., Santella, R.M., and Terry, M.B. (2012). Global DNA methylation levels in

white blood cell DNA from sisters discordant for breast cancer from the New York site of the Breast Cancer Family Registry. *Epigenetics* 7, 868–874.

Denis, H., Ndlovu, M.N., and Fuks, F. (2011). Regulation of mammalian DNA methyltransferases: a route to new mechanisms. *EMBO Rep.* 12, 647–656.

Deroo, L.A., Bolick, S.C.E., Xu, Z., Umbach, D.M., Shore, D., Weinberg, C.R., Sandler, D.P., and Taylor, J.A. (2014). Global DNA methylation and one-carbon metabolism gene polymorphisms and the risk of breast cancer in the Sister Study. *Carcinogenesis* 35, 333–338.

De Stefani, E., Boffetta, P., Ronco, a L., Deneo-Pellegrini, H., Correa, P., Acosta, G., Mendilaharsu, M., Luaces, M.E., and Silva, C. (2012). Processed meat consumption and risk of cancer: a multisite case-control study in Uruguay. *Br. J. Cancer* 107, 1584–1588.

Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., et al. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* 41, 1350–1353.

Dolinoy, D.C. (2008). The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. *Nutr. Rev.* 66 Suppl 1, S7–11.

Dolinoy, D.C., Weidman, J.R., and Jirtle, R.L. (2007). Epigenetic gene regulation: linking early developmental environment to adult disease. *Reprod. Toxicol.* 23, 297–307.

Du, P., Kibbe, W.A., and Lin, S.M. (2008). lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 24, 1547–1548.

Duffy, D., Zhao, Z., Sturm, R., Hayward, N., Martin, N., and Montgomery, G. (2013). Multiple pigmentation gene polymorphisms account for a substantial proportion of risk of cutaneous malignant melanoma. *J Invest Dermatol* 130, 520–528.

Dumitrescu, R.G., and Cotarla, I. (2005). Understanding breast cancer risk -- where do we stand in 2005? J. Cell. Mol. Med. 9, 208–221.

Duncan, B.K., and Miller, J.H. (1980). Mutagenic deamination of cytosine residues in DNA. *Nature* 287, 560–561.

Easton, D.F., Pooley, K.A., Dunning, A.M., Pharoah, P.D.P., Thompson, D., Ballinger, D.G., Struewing, J.P., Morrison, J., Field, H., Luben, R., et al. (2007). Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447, 1087–1093.

Ecsedi, S.I., Hernandez-Vargas, H., Lima, S.C., Herceg, Z., Adany, R., and Balazs, M. (2013). Transposable hypomethylation is associated with metastatic capacity of primary melanomas. *Int. J. Clin. Exp. Pathol.* 6, 2943–2948.

Eden, A., Gaudet, F., Waghmare, A., and Jaenisch, R. (2003). Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* 300, 455.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210.

Eggermont, A.M., Spatz, A., and Robert, C. (2013). Cutaneous melanoma. *Lancet* 6736, 1–12.

ENCODE Project Consortium. (2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 306, 636–640.

Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* 28, 817–825.

Esteller, M., Toyota, M., Sanchez-Cespedes, M., Capella, G., Peinado, M.A., Watkins, D.N., Issa, J.P., Sidransky, D., Baylin, S.B., and Herman, J.G. (2000). Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is associated with G to A mutations in K-ras in colorectal tumorigenesis. *Cancer Res.* 60, 2368–2371.

Esteller, M., Fraga, M.F., Guo, M., Garcia-Foncillas, J., Hedenfalk, I., Godwin, A.K., Trojan, J., Vaurs-Barrière, C., Bignon, Y.J., Ramus, S., et al. (2001). DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum. Mol. Genet.* 10, 3001–3007.

Esteller, M. (2008). Epigenetics in cancer. N. Engl. J. Med. 358, 1148-1159.

Eto, N., Miyagishi, M., Inagi, R., Fujita, T., and Nangaku, M. (2009). Mitogen-activated protein 3 kinase 6 mediates angiogenic and tumorigenic effects via vascular endothelial growth factor expression. *Am. J. Pathol.* 174, 1553–1563.

Falush, D., Stephens, M., and Pritchard, J.K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* 7, 574–578.

Fang, S., Han, J., Zhang, M., Wang, L.-E., Wei, Q., Amos, C.I., and Lee, J.E. (2013). Joint effect of multiple common SNPs predicts melanoma susceptibility. *PLoS One* 8, e85642.

Feil, R. (2006). Environmental and nutritional effects on the epigenetic regulation of genes. *Mutat. Res.* 600, 46–57.

Feinberg, A.P., and Irizarry, R.A. (2010). Evolution in health and medicine Sackler colloquium: Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl. Acad. Sci. U. S. A.* 107 Suppl , 1757–1764.

Fejerman, L., John, E.M., Huntsman, S., Beckman, K., Choudhry, S., Perez-Stable, E., Burchard, E.G., and Ziv, E. (2008). Genetic ancestry and risk of breast cancer among U.S. Latinas. *Cancer Res.* 68, 9723–9728.

Fejerman, L., Romieu, I., John, E.M., Lazcano-Ponce, E., Huntsman, S., Beckman, K.B., Pérez-Stable, E.J., González Burchard, E., Ziv, E., and Torres-Mejía, G. (2010). European ancestry is positively associated with breast cancer risk in Mexican women. *Cancer Epidemiol. Biomarkers Prev.* 19, 1074–1082.

Fejerman, L., Chen, G.K., Eng, C., Huntsman, S., Hu, D., Williams, A., Pasaniuc, B., John, E.M., Via, M., Gignoux, C., et al. (2012). Admixture mapping identifies a locus on 6q25 associated with breast cancer risk in US Latinas. *Hum. Mol. Genet.* 21, 1907–1917.

Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C., and Parkin, D.M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* 127, 2893–2917.

Fernandez, A.F., Assenov, Y., Martin-Subero, J.I., Balint, B., Siebert, R., Taniguchi, H., Yamamoto, H., Hidalgo, M., Tan, A.-C., Galm, O., et al. (2011). A DNA methylation fingerprint of 1628 human samples. *Genome Res.* 22, 407–419.

Figueiredo, J.C., Grau, M. V, Wallace, K., Levine, A.J., Shen, L., Hamdan, R., Chen, X., Bresalier, R.S., McKeown-Eyssen, G., Haile, R.W., et al. (2009). Global DNA hypomethylation (LINE-1) in the normal colon and lifestyle characteristics and dietary and genetic factors. *Cancer Epidemiol. Biomarkers Prev.* 18, 1041–1049.

Flanagan, J.M., Munoz-Alegre, M., Henderson, S., Tang, T., Sun, P., Johnson, N., Fletcher, O., Dos Santos Silva, I., Peto, J., Boshoff, C., et al. (2009). Gene-body hypermethylation of ATM in peripheral blood DNA of bilateral breast cancer patients. *Hum. Mol. Genet.* 18, 1332–1342.

Flanagan, J.M., Cocciardi, S., Waddell, N., Johnstone, C.N., Marsh, A., Henderson, S., Simpson, P., da Silva, L., Khanna, K., Lakhani, S., et al. (2010). DNA methylome of familial breast cancer identifies distinct profiles defined by mutation status. *Am. J. Hum. Genet.* 86, 420–433.

Forbes, S.A., Bindal, N., Bamford, S., Cole, C., Kok, C.Y., Beare, D., Jia, M., Shepherd, R., Leung, K., Menzies, A., et al. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* 39, D945–50.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suñer, D., Cigudosa, J.C., Urioste, M., Benitez, J., et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. U. S. A.* 102, 10604–10609.

Franchina, M., and Kay, P.H. (2001). Novel nucleotide substitutions within the coding region of DNMT2 are in strong linkage disequilibrium in Caucasians and Japanese. *Hum. Hered.* 52, 210–216.

Fraser, H.B., Lam, L.L., Neumann, S.M., and Kobor, M.S. (2012). Population-specificity of human DNA methylation. *Genome Biol.* 13, R8.

Friedman, R.C., Farh, K.K.-H., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19, 92–105.

Frigola, J., Ribas, M., Risques, R.A., and Peinado, M.A (2002). Methylome profiling of cancer cells by amplification of inter-methylated sites (AIMS). *Nucleic Acids Res.* 30, e28.

Frigola, J., Muñoz, M., Clark, S.J., Moreno, V., Capellà, G., and Peinado, M.A (2005a). Hypermethylation of the prostacyclin synthase (PTGIS) promoter is a frequent event in colorectal cancer and associated with aneuploidy. *Oncogene* 24, 7320–7326.

Frigola, J., Sole, X., Paz, M.F., Moreno, V., Esteller, M., Capella, G., and Peinado, M.A. (2005b). Differential DNA hypermethylation and hypomethylation signatures in colorectal cancer. *Hum. Mol. Genet.* 14, 319–326.

Frigola, J., Song, J., Stirzaker, C., Hinshelwood, R.A., Peinado, M.A., and Clark, S.J. (2006). Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. *Nat. Genet.* 38, 540–549.

Friso, S., Udali, S., Guarini, P., Pellegrini, C., Pattini, P., Moruzzi, S., Girelli, D., Pizzolo, F., Martinelli, N., Corrocher, R., et al. (2013). Global DNA hypomethylation in peripheral blood mononuclear cells as a biomarker of cancer risk. *Cancer Epidemiol. Biomarkers Prev.* 22, 348–355.

Fuke, C., Shimabukuro, M., Petronis, A., Sugimoto, J., Oda, T., Miura, K., Miyazaki, T., Ogura, C., Okazaki, Y., and Jinno, Y. (2004). Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study. *Ann. Hum. Genet.* 68, 196–204.

Gandini, S., Sera, F., Cattaruzza, M.S., Pasquini, P., Zanetti, R., Masini, C., Boyle, P., and Melchi, C.F. (2005). Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors. *Eur. J. Cancer* 41, 2040–2059.

Gao, Y., Baccarelli, A., Shu, X.O., Ji, B.-T., Yu, K., Tarantini, L., Yang, G., Li, H.-L., Hou, L., Rothman, N., et al. (2012). Blood leukocyte Alu and LINE-1 methylation and gastric cancer risk in the Shanghai Women's Health Study. *Br. J. Cancer* 106, 585–591.

Garbe, C., and Leiter, U. (2009). Melanoma epidemiology and trends. *Clin. Dermatol.* 27, 3–9.

Garrick, D., Fiering, S., Martin, D.I., and Whitelaw, E. (1998). Repeat-induced gene silencing in mammals. *Nat. Genet.* 18, 56–59.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 5, R80.

Gibbs, J.R., van der Brug, M.P., Hernandez, D.G., Traynor, B.J., Nalls, M.A., Lai, S.-L., Arepalli, S., Dillman, A., Rafferty, I.P., Troncoso, J., et al. (2010). Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* 6, e1000952.

Gluckman, P.D., Hanson, M.A., and Spencer, H.G. (2005). Predictive adaptive responses and human evolution. *Trends Ecol. Evol.* 20, 527–533.

Gomes, M.V.M., Toffoli, L. V, Arruda, D.W., Soldera, L.M., Pelosi, G.G., Neves-Souza, R.D., Freitas, E.R., Castro, D.T., and Marquez, A.S. (2012). Age-related changes in the global DNA methylation profile of leukocytes are linked to nutrition but are not associated with the MTHFR C677T genotype or to functional capacities. *PLoS One* 7, e52570.

Greenberg, E.S., Chong, K.K., Huynh, K.T., Tanaka, R., and Hoon, D.S.B. (2012). Epigenetic biomarkers in skin cancer. *Cancer Lett.* 342, 170.

Guerrero-Preston, R., Santella, R., Desai, M., Berdasco, M., and Fraga, M. (2007). Global DNA Hypomethylation in Liver Cancer Cases and Controls. *Epigenetics* 2, 223–226.

Guibert, S., and Weber, M. (2013). Functions of DNA methylation and hydroxymethylation in mammalian development. In Current Topics in Developmental Biology, (2013 Elsevier Inc.), pp. 47–83.

Hackett, J.A., and Surani, M.A. (2013). DNA methylation dynamics during the mammalian life cycle. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 368, 20110328.

Hales, C.N., Barker, D.J., Clark, P.M., Cox, L.J., Fall, C., Osmond, C., and Winter, P.D. (1991). Fetal and infant growth and impaired glucose tolerance at age 64. *BMJ* 303, 1019–1022.

Hanson, M. a, Low, F.M., and Gluckman, P.D. (2011). Epigenetic epidemiology: the rebirth of soft inheritance. *Ann. Nutr. Metab.* 58 Suppl 2, 8–15.

He, L., and Hannon, G.J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* 5, 522–531.

He, X.-J., Chen, T., and Zhu, J.-K. (2011). Regulation and function of DNA methylation in plants and animals. *Cell Res.* 21, 442–465.

Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G.J., Susser, E.S., Slagboom, P.E., and Lumey, L.H. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci. U. S. A.* 105, 17046–17049.

Herceg, Z., and Vaissière, T. (2011). Epigenetic mechanisms and cancer: An interface between the environment and the genome. *Epigenetics* 6, 804–819.

Herman, J.G., and Baylin, S.B. (2003). Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.* 349, 2042–2054.

Heyn, H., Moran, S., Hernando-Herraez, I., Sayols, S., Gomez, A., Sandoval, J., Monk, D., Hata, K., Marques-Bonet, T., Wang, L., et al. (2013a). DNA methylation contributes to natural human variation. *Genome Res.* 23, 1363–1372.

Heyn, H., Carmona, F.J., Gomez, A., Ferreira, H.J., Bell, J.T., Sayols, S., Ward, K., Stefansson, O.A., Moran, S., Sandoval, J., et al. (2013b). DNA methylation profiling in breast cancer discordant identical twins identifies DOK7 as novel epigenetic biomarker. *Carcinogenesis* 34, 102–108.

Hicks, C., Kumar, R., Pannuti, A., Backus, K., Brown, A., Monico, J., and Miele, L. (2013). An Integrative Genomics Approach for Associating GWAS Information with Triple-Negative Breast Cancer. *Cancer Inform.* 12, 1–20.

Hidalgo, P.C., Bengochea, M., Abilleira, D., Cabrera, A., and Alvarez, I. (2005). Genetic admixture estimate in the Uruguayan population based on the loci LDLR, GYPA, HBGG, GC and D7S8. *Int. J. Hum. Genet.* 5, 217–222.

High, W.A., and Robinson, W.A. (2007). Genetic mutations involved in melanoma: a summary of our current understanding. *Adv. Dermatol.* 23, 61–79.

Hitchins, M.P. (2010). Inheritance of Epigenetic Aberrations (Constitutional Epimutations) in Cancer Susceptibility. In Epigenetics and Cancer, Part A, (Elsevier Inc.), pp. 201–243.

Hitchins, M.P., Wong, J.J.L., Suthers, G., Suter, C.M., Martin, D.I.K., Hawkins, N.J., and Ward, R. (2007). Inheritance of a cancer-associated MLH1 germ-line epimutation. *N. Engl. J. Med.* 356, 697–705.

Hochmann, J. (2012). Identificación y caracterización de regiones cromosómicas asociadas a melanoma esporádico en el Uruguay mediante el método de mapeo por mestizaje: MC1R como posible gen candidato. Tesis de Maestría, PEDECIBA -Universidad de la República.

Hochmann, J., Cappetta, M., Pérez, J., Colistro, V., Larre Borges, A., Nicolletti, S., Velázquez, T., Souto, J., Roche, L., Rivas, G., et al. (2010). Identificación de regiones genómicas asociadas a melanoma esporádico en Uruguay. In XIV Congreso Latinoamericano de Genética (ALAG2010).

Holliday, R. (1990). Mechanisms for the control of gene activity during development. *Biol. Rev. Camb. Philos. Soc.* 65, 431–471.

Holm, K., Hegardt, C., Staaf, J., Vallon-Christersson, J., Jönsson, G., Olsson, H., Borg, A., and Ringnér, M. (2010). Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns. *Breast Cancer Res.* 12, R36.

Hon, G.C., Hawkins, R.D., Caballero, O.L., Lo, C., Lister, R., Pelizzola, M., Valsesia, A., Ye, Z., Kuan, S., Edsall, L.E., et al. (2012). Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res.* 22, 246–258.

Hou, L., Wang, H., Sartori, S., Gawron, A., Lissowska, J., Bollati, V., Tarantini, L., Zhang, F.F., Zatonski, W., and Baccarelli, A. (2010). Blood DNA hypomethylation and gastric cancer risk in a high-risk Polish population. *Int. J. Cancer* 127, 1866–1874.

Hsiung, D.T., Marsit, C.J., Houseman, E.A., Eddy, K., Furniss, C.S., McClean, M.D., and Kelsey, K.T. (2007). Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. *Cancer Epidemiol. Biomarkers Prev.* 16, 108–114.

Hu, Z., Fan, C., Oh, D.S., Marron, J.S., He, X., Qaqish, B.F., Livasy, C., Carey, L.A., Reynolds, E., Dressler, L., et al. (2006). The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* 7, 96.

Huang, D.W., Sherman, B.T., Zheng, X., Yang, J., Imamichi, T., Stephens, R., and Lempicki, R.A. (2009). Extracting biological meaning from large gene lists with DAVID. *Curr. Protoc. Bioinformatics* Chapter 13, Unit 13.11.

Huang, Y., Nayak, S., Jankowitz, R., Davidson, N.E., and Oesterreich, S. (2011). Epigenetics in breast cancer : what 's new? *Breast Cancer Res.* 13, 225.

Huang, Y.-W., Huang, T.H.-M., and Wang, L.-S. (2010). Profiling DNA methylomes from microarray to genome-scale sequencing. *Technol. Cancer Res. Treat.* 9, 139–147.

Huidobro, C., Fernandez, A.F., and Fraga, M.F. (2013). The role of genetics in the establishment and maintenance of the epigenome. *Cell. Mol. Life Sci.* 70, 1543–1573.

Hyland, P.L., Burke, L.S., Pfeiffer, R.M., Mirabello, L., Tucker, M. a, Goldstein, A.M., and Yang, X.R. (2013). LINE-1 methylation in peripheral blood and the risk of melanoma in melanoma-prone families with and without CDKN2A mutations. *Melanoma Res.* 23, 55–60.

Ichiyanagi, K. (2013). Epigenetic regulation of transcription and possible functions of mammalian short interspersed elements, SINEs. *Genes Genet. Syst.* 88, 19–29.

Inbar-Feigenberg, M., Choufani, S., Butcher, D.T., Roifman, M., and Weksberg, R. (2013). Basic concepts of epigenetics. *Fertil. Steril.* 99, 607–615.

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.* 41, 178–186.

Ito, Y., Koessler, T., Ibrahim, A.E.K., Rai, S., Vowler, S.L., Abu-Amero, S., Silva, A.-L., Maia, A.-T., Huddleston, J.E., Uribe-Lewis, S., et al. (2008). Somatically acquired hypomethylation of IGF2 in breast and colorectal cancer. *Hum. Mol. Genet.* 17, 2633–2643.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* 33 Suppl, 245–254.

James, S.J., Pogribny, I.P., Pogribna, M., Miller, B.J., Jernigan, S., and Melnyk, S. (2003). Mechanisms of DNA damage, DNA hypomethylation, and tumor progression in the folate/methyl-deficient rat model of hepatocarcinogenesis. *J. Nutr.* 133, 3740S–3747S.

Javierre, B.M., Fernandez, A.F., Richter, J., Al-Shahrour, F., Martin-Subero, J.I., Rodriguez-Ubreva, J., Berdasco, M., Fraga, M.F., O'Hanlon, T.P., Rider, L.G., et al. (2010). Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res.* 20, 170–179.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* 293, 1074–1080.

Jin, B., and Robertson, K.D. (2013). DNA methyltransferases, DNA damage repair, and cancer. *Adv. Exp. Med. Biol.* 754, 3–29.

Jirtle, R.L., and Skinner, M.K. (2007). Environmental epigenomics and disease susceptibility. *Nat. Rev. Genet.* 8, 253–262.

Johnson, L.J., and Tricker, P.J. (2010). Epigenomic plasticity within populations: its evolutionary significance and potential. *Heredity* (Edinb). 105, 113–121.

Jordá, M., Rodriguez, J., Frigola, J., and Peinado, M.A. (2009). Analysis of DNA methylation by amplification of intermethylated sites (AIMS). In DNA Methylation: Methods and Protocols, J. Tost, ed. (Humana Press), pp. 107–116.

Jovanovic, J., Rønneberg, J.A., Tost, J., and Kristensen, V. (2010). The epigenetics of breast cancer. *Mol. Oncol.* 4, 242–254.

Kaikkonen, M.U., Lam, M.T.Y., and Glass, C.K. (2011). Non-coding RNAs as regulators of gene expression and epigenetics. *Cardiovasc. Res.* 90, 430–440.

Kaminsky, Z.A., Tang, T., Wang, S.-C., Ptak, C., Oh, G.H.T., Wong, A.H.C., Feldcamp, L.A., Virtanen, C., Halfvarson, J., Tysk, C., et al. (2009). DNA methylation profiles in monozygotic and dizygotic twins. *Nat. Genet.* 41, 240–245.

Katari, S., Turan, N., Bibikova, M., Erinle, O., Chalian, R., Foster, M., Gaughan, J.P., Coutifaris, C., and Sapienza, C. (2009). DNA methylation and gene expression differences in children conceived in vitro or in vivo. *Hum. Mol. Genet.* 18, 3769–3778.

Katto, J., and Mahlknecht, U. (2011). Epigenetic regulation of cellular adhesion in cancer. *Carcinogenesis* 32, 1414–1418.

Khan, S.I., Aumsuwan, P., Khan, I. a, Walker, L. a, and Dasmahapatra, A.K. (2012). Epigenetic events associated with breast cancer and their prevention by dietary components targeting the epigenome. *Chem. Res. Toxicol.* 25, 61–73.

Khoja, L., Lorigan, P., Zhou, C., Lancashire, M., Booth, J., Cummings, J., Califano, R., Clack, G., Hughes, A., and Dive, C. (2013). Biomarker utility of circulating tumor cells in metastatic cutaneous melanoma. *J. Invest. Dermatol.* 133, 1582–1590.

Kok, R.M., Smith, D.E.C., Barto, R., Spijkerman, A.M.W., Teerlink, T., Gellekink, H.J., Jakobs, C., and Smulders, Y.M. (2007). Global DNA methylation measured by liquid chromatography-tandem mass spectrometry: analytical technique, reference values and determinants in healthy subjects. *Clin. Chem. Lab. Med.* 45, 903–911.

Korshunova, Y., Maloney, R.K., Lakey, N., Citek, R.W., Bacher, B., Budiman, A., Ordway, J.M., McCombie, W.R., Leon, J., Jeddeloh, J.A., et al. (2008). Massively parallel bisulphite pyrosequencing reveals the molecular complexity of breast cancer-associated cytosine-methylation patterns obtained from tissue and serum DNA. *Genome Res.* 18, 19–29.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell 128, 693–705.

Koyanagi, K., Mori, T., Day, S.J.O., Martinez, S.R., Wang, H., and Hoon, D.S.B. (2010). Association of Circulating Tumor Cells with Serum Tumor- Related Methylated DNA in Peripheral Blood of Melanoma Patients. *Cancer Res.* 66, 6111–6117.

Krizsan-agbas, D., Pedchenko, T., and Smith, P.G. (2009). Neurotrimin is an estrogenregulated determinant of peripheral sympathetic innervation. *J Neurosci Res* 86, 3086–3095.

Kulis, M., Queirós, A.C., Beekman, R., and Martín-Subero, J.I. (2013). Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer. *Biochim. Biophys. Acta* 1829, 1161–1174.

Kupershmidt, I., Su, Q.J., Grewal, A., Sundaresh, S., Halperin, I., Flynn, J., Shekar, M., Wang, H., Park, J., Cui, W., et al. (2010). Ontology-based meta-analysis of global collections of high-throughput public data. *PLoS One* 5, e13066.

Kwabi-Addo, B., Wang, S., Chung, W., Jelinek, J., Patierno, S.R., Wang, B.-D., Andrawis, R., Lee, N.H., Apprey, V., Issa, J.-P., et al. (2010). Identification of differentially methylated genes in normal prostate tissues from African American and Caucasian men. *Clin. Cancer Res.* 16, 3539–3547.

LaBreche, H.G., Nevins, J.R., and Huang, E. (2011). Integrating factor analysis and a transgenic mouse model to reveal a peripheral blood predictor of breast tumors. *BMC Med. Genomics* 4, 61.

Laird, P.W. (2010). Principles and challenges of genomewide DNA methylation analysis. *Nat. Rev. Genet.* 11, 191–203.

Lander, E.S. (2011). Initial impact of the sequencing of the human genome. *Nature* 470, 187–197.

Lange, C.P.E., Campan, M., Hinoue, T., Schmitz, R.F., van der Meulen-de Jong, A.E., Slingerland, H., Kok, P.J.M.J., van Dijk, C.M., Weisenberger, D.J., Shen, H., et al. (2012). Genome-scale discovery of DNA-methylation biomarkers for blood-based detection of colorectal cancer. *PLoS One* 7, e50266.

Larre Borges, A., Borges, A.L., Cuéllar, F., Puig-Butillé, J.A., Scarone, M., Delgado, L., Badenas, C., Milà, M., Malvehy, J., Barquet, V., et al. (2009). CDKN2A mutations in melanoma families from Uruguay. *Br. J. Dermatol.* 161, 536–541.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Kin Sung, K.W., Rigoutsos, I., Loring, J., et al. (2010). Dynamic changes in the human methylome during differentiation. *Genome Res.* 20, 320–331.

Lee, B., Morano, A., Porcellini, A., and Muller, M.T. (2012). GADD45α inhibition of DNMT1 dependent DNA methylation during homology directed DNA repair. *Nucleic Acids Res.* 40, 2481–2493.

Lee, J., Jang, S.J., Benoit, N., Hoque, M.O., Califano, J.A., Trink, B., Sidransky, D., Mao, L., and Moon, C. (2010). Presence of 5-methylcytosine in CpNpG trinucleotides in the human genome. *Genomics* 96, 67–72.

Li, B., Carey, M., and Workman, J.L. (2007). The Role of chromatin during transcription. *Cell* 128, 707–719.

Li, J.L., Mazar, J., Zhong, C., Faulkner, G.J., Govindarajan, S.S., Zhang, Z., Dinger, M.E., Meredith, G., Adams, C., Zhang, S., et al. (2013). Genome-wide methylated CpG island profiles of melanoma cells reveal a melanoma coregulation network. *Sci. Rep.* 3, 2962.

Li, L., Choi, J.-Y., Lee, K.-M., Sung, H., Park, S.K., Oze, I., Pan, K.-F., You, W.-C., Chen, Y.-X., Fang, J.-Y., et al. (2012). DNA methylation in peripheral blood: a potential biomarker for cancer molecular epidemiology. J. Epidemiol. 22, 384–394.

Lian, C.G., Xu, Y., Ceol, C., Wu, F., Larson, A., Xu, W., Tan, L., Hu, Y., Zhan, Q., Lee, C., et al. (2012). Loss of 5-hidroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell* 150, 1135–1146.

Lim, U., Flood, A., Choi, S.-W., Albanes, D., Cross, A.J., Schatzkin, A., Sinha, R., Katki, H. a, Cash, B., Schoenfeld, P., et al. (2008). Genomic methylation of leukocyte DNA in relation to colorectal adenoma among asymptomatic women. *Gastroenterology* 134, 47–55.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462, 315–322.

Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., Antosiewicz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., et al. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* 471, 68–73.

Liu, H., Zhou, Y., Boggs, S.E., Belinsky, S.A., and Liu, J. (2007). Cigarette smoke induces demethylation of prometastatic oncogene synuclein-gamma in lung cancer cells by downregulation of DNMT3B. *Oncogene* 26, 5900–5910.

Lo, P.-K., and Sukumar, S. (2008). Epigenomics and breast cancer. *Pharmacogenomics* 9, 1879–1902.

Loizidou, M.A., Cariolou, M.A., Neuhausen, S.L., Newbold, R.F., Bashiardes, E., Marcou, Y., Michael, T., Daniel, M., Kakouri, E., Papadopoulos, P., et al. (2010). Genetic variation in genes interacting with BRCA1/2 and risk of breast cancer in the Cypriot population. *Breast Cancer Res. Treat.* 121, 147–156.

Lopez-Serra, P., and Esteller, M. (2012). DNA methylation-associated silencing of tumorsuppressor microRNAs in cancer. *Oncogene* 31, 1609–1622.

Luciani, S., Cabanes, A., Prieto-Lara, E., and Gawryszewski, V. (2013). Cervical and female breast cancers in the Americas: current situation and opportunities for action. *Bull. World Health Organ.* 91, 640–649.

Luiz, O.C., Gianini, R.J., Gonçalves, F.T., Francisco, G., Festa-Neto, C., Sanches, J.A., Gattas, G.J.F., Chammas, R., and Eluf-Neto, J. (2012). Ethnicity and cutaneous melanoma in the city of Sao Paulo, Brazil: a case-control study. *PLoS One* 7, e36348.

Lujambio, A., Calin, G.A., Villanueva, A., Ropero, S., Sánchez-Céspedes, M., Blanco, D., Montuenga, L.M., Rossi, S., Nicoloso, M.S., Faller, W.J., et al. (2008). A microRNA DNA methylation signature for human cancer metastasis. *PNAS* 105, 13556–13561.

Ma, H., Zhang, W., Hu, J., Yu, Z., Chen, Y., Luo, Q., Shi, X., Zhang, Y., Song, R., Zhou, Z., et al. (2009). Analysis of global DNA methylation levels in human blood using high-performance liquid chromatography/tandem electrospray ionization mass spectrometry. *Eur. J. Mass Spectrom.* 15, 555–561.

Maegawa, S., Hinkal, G., Kim, H.S., Shen, L., Zhang, L., Zhang, J., Zhang, N., Liang, S., Donehower, L.A., and Issa, J.P. (2010). Widespread and tissue specific age-related DNA methylation changes in mice. *Genome Res.* 20, 332–340.

Marsit, C.J., Koestler, D.C., Christensen, B.C., Karagas, M.R., Houseman, E.A., and Kelsey, K.T. (2011). DNA methylation array analysis identifies profiles of blood-derived DNA methylation associated with bladder cancer. *J. Clin. Oncol.* 29, 1133–1139.

Mazzei, E., Hochmann, J., Manrique, G., Mariño, A.L., Delgado, L., and Asuaga, M.M. (2013). Determinación de la mutación BRAF V600E en melanomas de pacientes uruguayos. *Rev. Médica del Uruguay* 29, 97–102.

McGowan, P.O., Sasaki, A., D'Alessio, A.C., Dymov, S., Labonté, B., Szyf, M., Turecki, G., and Meaney, M.J. (2009). Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nat. Neurosci.* 12, 342–348.

McNamee, C.J., Reed, J.E., Howard, M.R., Lodge, A.P., and Moss, D.J. (2002). Promotion of neuronal cell adhesion by members of the IgLON family occurs in the absence of either support or modification of neurite outgrowth. *J. Neurochem.* 80, 941–948.

McPherson, J.P., Lemmers, B., Chahwan, R., Pamidi, A., Migon, E., Matysiak-Zablocki, E., Moynahan, M.E., Essers, J., Hanada, K., Poonepalli, A., et al. (2004). Involvement of mammalian Mus81 in genome integrity and tumor suppression. *Science* 304, 1822–1826.

Mohammed, S.I., Springfield, S., and Das, R. (2012). Role of epigenetics in cancer health disparities. In Methods in Molecular Biology (Clifton, N.J.), pp. 395–410.

Moore, L.E., Pfeiffer, R.M., Poscablo, C., Real, F.X., Kogevinas, M., Silverman, D., García-Closas, R., Chanock, S., Tardón, A., Serra, C., et al. (2008). Genomic DNA hypomethylation as a biomarker for bladder cancer susceptibility in the Spanish Bladder Cancer Study: a case-control study. *Lancet Oncol.* 9, 359–366.

Mosca, E., Alfieri, R., Merelli, I., Viti, F., Calabria, A., and Milanesi, L. (2010). A multilevel data integration resource for breast cancer study. *BMC Syst. Biol.* 4, 76.

Mulero-Navarro, S., and Esteller, M. (2008). Epigenetic biomarkers for human cancer: the time is now. *Crit. Rev. Oncol. Hematol.* 68, 1–11.

Muthusamy, V., Duraisamy, S., Bradbury, C.M., Hobbs, C., Curley, D.P., Nelson, B., and Bosenberg, M. (2006). Epigenetic silencing of novel tumor suppressors in malignant melanoma. *Cancer Res.* 66, 11187–11193.

Nakken, S., Alseth, I., and Rognes, T. (2007). Computational prediction of the effects of non-synonymous single nucleotide polymorphisms in human DNA repair genes. *Neuroscience* 145, 1273–1279.

Narayanan, S., McConnell, J., Little, J., Sharp, L., Piyathilake, C.J., Powers, H., Basten, G., and Duthie, S.J. (2004). Associations between two common variants C677T and A1298C in the methylenetetrahydrofolate reductase gene and measures of folate metabolism and DNA stability (strand breaks, misincorporated uracil, and DNA methylation status) in human lymphocytes in vivo. *Cancer Epidemiol. Biomarkers Prev.* 13, 1436–1443.

Ntougkos, E., Rush, R., Scott, D., Frankenberg, T., Gabra, H., Smyth, J.F., and Sellar, G.C. (2005). The IgLON family in epithelial ovarian cancer: expression profiles and clinicopathologic correlates. *Clin. Cancer Res.* 11, 5764–5768.

O'Reilly, K.E., de Miera, E.V.-S., Segura, M.F., Friedman, E., Poliseno, L., Han, S.W., Zhong, J., Zavadil, J., Pavlick, A., Hernando, E., et al. (2013). Hedgehog pathway blockade inhibits melanoma cell growth in vitro and in vivo. *Pharmaceuticals* 6, 1429–1450.

Osman, F., and Whitby, M.C. (2007). Exploring the roles of Mus81-Eme1/Mms4 at perturbed replication forks. *DNA Repair* 6, 1004–1017.

Otero, A.M., Pou Ferrari, R., Pons, E., Lens, D., De Lisa, E., Dellepiane, M., Storch, E., Attarian, D., Ferrari, A., Pierri, S., et al. (2004). Trombofilia y pérdida recurrente de embarazo. Rev. Médica Del Uruguay 20, 106–113.

Pan, Y., Wang, K.-S., and Aragam, N. (2011). NTM and NR3C2 polymorphisms influencing intelligence: family-based association studies. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 35, 154–160.

Paz, M.F., Avila, S., Fraga, M.F., Pollan, M., Capella, G., Peinado, M.A., Sanchez-Cespedes, M., Herman, J.G., and Esteller, M. (2002). Germ-line variants in methyl-group metabolism genes and susceptibility to DNA methylation in normal tissues and human primary tumors. *Cancer Res.* 62, 4519–4524.

Paz, M.F., Wei, S., Cigudosa, J.C., Rodriguez-Perales, S., Peinado, M.A., Huang, T.H.-M., and Esteller, M. (2003). Genetic unmasking of epigenetically silenced tumor suppressor genes in colon cancer cells deficient in DNA methyltransferases. *Hum. Mol. Genet.* 12, 2209–2219.

Pedersen, K.S., Bamlet, W.R., Oberg, A.L., De Andrade, M., Matsumoto, M.E., Tang, H., Thibodeau, S.N., Petersen, G.M., and Wang, L. (2011). Leukocyte DNA methylation signature differentiates pancreatic cancer patients from healthy controls. *PLoS One* 6, 9.

Peto, J. (2001). Cancer epidemiology in the last century and the next decade. *Nature* 411, 390–395.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. *Nat. Biotechnol.* 28, 1057–1068.

Pufulete, M., Al-Ghnaniem, R., Leather, A.J.M., Appleby, P., Gout, S., Terry, C., Emery, P.W., and Sanders, T.A.B. (2003). Folate status, genomic DNA hypomethylation, and risk of colorectal adenoma and cancer: a case control study. *Gastroenterology* 124, 1240–1248.

R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/. R Found. Stat. Comput. Vienna, Austria.

Rakyan, V.K., Down, T.A., Balding, D.J., and Beck, S. (2012). Epigenome-Wide Association Studies for common human diseases. *Nat. Rev. Genet.* 12, 529–541.

Raskin, L., Fullen, D.R., Giordano, T.J., Thomas, D.G., Frohm, M.L., Cha, K.B., Ahn, J., Mukherjee, B., Johnson, T.M., and Gruber, S.B. (2013). Transcriptome profiling identifies HMGA2 as a biomarker of melanoma progression and prognosis. *J. Invest. Dermatol.* 133, 2585–2592.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science* 293, 1089–1093.

Richardson, A.L., Wang, Z.C., De Nicolo, A., Lu, X., Brown, M., Miron, A., Liao, X., Iglehart, J.D., Livingston, D.M., and Ganesan, S. (2006). X chromosomal abnormalities in basal-like human breast cancer. *Cancer Cell* 9, 121–132.

Rideout, W.M., Coetzee, G.A., Olumi, A.F., and Jones, P.A. (1990). 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* 249, 1288–1290.

Risch, A., Sarkisyan, N., Scherf, D., Jacobsson, H., Hagmann, W., and Plass, C. (2012). Epigenetic epidemiology of cancer. In Epigenetic Epidemiology, K.B. Michels, ed. (Dordrecht: Springer Netherlands), pp. 225–267.

Rivera, C.M., and Ren, B. (2013). Mapping human epigenomes. Cell 155, 39-55.

Robertson, K.D., and Wolffe, A.P. (2000). DNA methylation in health and disease. *Nat. Rev. Genet.* 1, 11–19.

Rodriguez, J., Frigola, J., Vendrell, E., Risques, R.-A., Fraga, M.F., Morales, C., Moreno, V., Esteller, M., Capellà, G., Ribas, M., et al. (2006). Chromosomal instability correlates with genome-wide DNA demethylation in human primary colorectal cancers. *Cancer Res.* 66, 8462–9468.

Roff, D.A., and Bentzen, P. (1989). The statistical analysis of mitochondrial DNA polymorphisms: chi 2 and the problem of small samples. *Mol. Biol. Evol.* 6, 539–545.

Ronco, A.L., de Stefani, E., Aune, D., Boffetta, P., Deneo-Pellegrini, H., Acosta, G., and Mendilaharsu, M. (2010). Nutrient patterns and risk of breast cancer in Uruguay. *Asian Pac. J. Cancer Prev.* 11, 519–524.

Ronco, A.L., Stefani, E. De, Correa, P., Deneo-Pellegrini, H., Boffetta, P., Acosta, G., and Mendilaharsu, M. (2011). Dietary Benzo[a]pyrene, Alcohol Drinking, and Risk of Breast Cancer: a Case-control Study in Uruguay. *Asian Pac. J. Cancer Prev.* 12, 1463–1467.

Ronco, A.L., De Stefani, E., and Deneo-Pellegrini, H. (2012). Risk factors for premenopausal breast cancer: a case-control study in Uruguay. *Asian Pacific J. Cancer Prev.* 13, 2879–2886.

Sadikovic, B., Haines, T.R., Butcher, D.T., and Rodenhiser, D.I. (2004). Chemically induced DNA hypomethylation in breast carcinoma cells detected by the amplification of intermethylated sites. *Breast Cancer Res.* 6, R329–37.

Sadikovic, B., Al-Romaih, K., Squire, J.A., and Zielenska, M. (2008). Cause and consequences of genetic and epigenetic alterations in human cancer. *Curr. Genomics* 9, 394–408.

Sandoval, J., Heyn, H., Moran, S., Serra-Musach, J., Pujana, M.A., Bibikova, M., and Esteller, M. (2011). Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* 6, 692–702.

Sans, M., Salzano, F., and Chakraborty, R. (1997). Historical Genetics in Uruguay: Estimates of Biological Origins and Their Problems. *Hum. Biol.* 69.

Schmitz, R.J., Schultz, M.D., Urich, M. a., Nery, J.R., Pelizzola, M., Libiger, O., Alix, A., McCosh, R.B., Chen, H., Schork, N.J., et al. (2013). Patterns of population epigenomic diversity. *Nature* 495, 193.

Sellar, G.C., Watt, K.P., Rabiasz, G.J., Stronach, E.A., Li, L., Miller, E.P., Massie, C.E., Miller, J., Contreras-Moreira, B., Scott, D., et al. (2003). OPCML at 11q25 is epigenetically inactivated and has tumor-suppressor function in epithelial ovarian cancer. *Nat. Genet.* 34, 337–343.

Sharma, S., Kelly, T.K., and Jones, P.A. (2010). Epigenetics in cancer. *Carcinogenesis* 31, 27–36.

Shenker, N., and Flanagan, J.M. (2012). Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research. *Br. J. Cancer* 106, 248–253.

Shukla, V., Coumoul, X., Lahusen, T., Wang, RH., Xu, X., Vassilopoulos, A., Xiao, C., Lee, MH., Man, YG., Ouchi, M., Ouchi, T., Deng, CX. (2010). BRCA1 affects global DNA methylation through regulation of DNMT1. *Cell Research* 20:1201-1215.

Sigalotti, L., Covre, A., Fratta, E., Parisi, G., Colizzi, F., Rizzo, A., Danielli, R., Nicolay, H.J.M., Coral, S., and Maio, M. (2010). Epigenetics of human cutaneous melanoma: setting the stage for new therapeutic strategies. *J. Transl. Med.* 8, 56.

Sigalotti, L., Fratta, E., Bidoli, E., Covre, A., Parisi, G., Colizzi, F., Coral, S., Massarut, S., Kirkwood, J.M., and Maio, M. (2011). Methylation levels of the "long interspersed nucleotide element-1" repetitive sequences predict survival of melanoma patients. *J. Transl. Med.* 9, 78.

Silva, J.M., Ezhkova, E., Silva, J., Heart, S., Castillo, M., Campos, Y., Castro, V., Bonilla, F., Cordon-Cardo, C., Muthuswamy, S.K., et al. (2009). Cyfip1 is a putative invasion suppressor in epithelial cancers. *Cell* 137, 1047–1061.

Skinner, M.K. (2011). Environmental epigenetic transgenerational inheritance and somatic epigenetic mitotic stability. *Epigenetics* 6, 838–842.

Slieker, R.C., Bos, S.D., Goeman, J.J., Bovée, J.V., Talens, R.P., van der Breggen, R., Suchiman, H.E.D., Lameijer, E.-W., Putter, H., van den Akker, E.B., et al. (2013). Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450K array. *Epigenetics Chromatin* 6, 26.

Smith, P., McGuffog, L., Easton, D.F., Mann, G.J., Pupo, G.M., Newman, B., Chenevix-Trench, G., Szabo, C., Southey, M., Renard, H., et al. (2006). A genome wide linkage search for breast cancer susceptibility genes. *Genes. Chromosomes Cancer* 45, 646–655.

Soares, J., Pinto, A.E., Cunha, C. V, André, S., Barão, I., Sousa, J.M., and Cravo, M. (1999). Global DNA hypomethylation in breast carcinoma: correlation with prognostic factors and tumor progression. *Cancer* 85, 112–118.

Sørlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., et al. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *PNAS* 98, 10869–10874.

Sproul, D., and Meehan, R.R. (2013). Genomic insights into cancer-associated aberrant CpG island hypermethylation. Brief. Funct. Genomics 12, 174–190.

Stefansson, O. A, and Esteller, M. (2013). Epigenetic modifications in breast cancer and their role in personalized medicine. *Am. J. Pathol.* 183, 1052–1063.

Steine, E.J., Ehrich, M., Bell, G.W., Raj, A., Reddy, S., van Oudenaarden, A., Jaenisch, R., and Linhart, H.G. (2011). Genes methylated by DNA methyltransferase 3b are similar in mouse intestine and human colon cancer. *J. Clin. Invest.* 121, 1748–1752.

Stephens, P.J., Tarpey, P.S., Davies, H., Van Loo, P., Greenman, C., Wedge, D.C., Nik-Zainal, S., Martin, S., Varela, I., Bignell, G.R., et al. (2012). The landscape of cancer genes and mutational processes in breast cancer. *Nature* 486, 400–404.

Stover, P.J. (2011). Polymorphisms in 1-carbon metabolism, epigenetics and folate-related pathologies. J. Nutrigenet. Nutrigenomics 4, 293–305.

Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., Simon, I., Yakhini, Z., and Cedar, H. (2009). Developmental programming of CpG island methylation profiles in the human genome. *Nat. Struct. Mol. Biol.* 16, 564–571.

Styrkarsdottir, U., Halldorsson, B. V, Gretarsdottir, S., Gudbjartsson, D.F., Walters, G.B., Ingvarsson, T., Jonsdottir, T., Saemundsdottir, J., Snorradóttir, S., Center, J.R., et al. (2009). New sequence variants associated with bone mineral density. *Nat. Genet.* 41, 15–17.
Suijkerbuijk, K.P.M., Fackler, M.J., Sukumar, S., van Gils, C.H., van Laar, T., van der Wall, E., Vooijs, M., and van Diest, P.J. (2008). Methylation is less abundant in BRCA1-associated compared with sporadic breast cancer. *Ann. Oncol.* 19, 1870–1874.

Suzuki, R., and Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22, 1540–1542.

Svingen, T., Beverdam, A., Bernard, P., McClive, P., Harley, V.R., Sinclair, A.H., and Koopman, P. (2007). Sex-specific expression of a novel gene Tmem184a during mouse testis differentiation. *Reproduction* 133, 983–989.

Taby, R., and Issa, J.-P.J. (2010). Cancer epigenetics. CA. Cancer J. Clin. 60, 376–392.

Takeuchi, T., Adachi, Y., and Ohtsuki, Y. (2005). Skeletrophin, a novel ubiquitin ligase to the intracellular region of Jagged-2, is aberrantly expressed in multiple myeloma. *Am. J. Pathol.* 166, 1817–1826.

Takeuchi, T., Adachi, Y., Sonobe, H., Furihata, M., and Ohtsuki, Y. (2006). A ubiquitin ligase, skeletrophin, is a negative regulator of melanoma invasion. *Oncogene* 25, 7059–7069.

Tanas, a. S., Shkarupo, V. V., Kuznetsova, E.B., Zaletayev, D. V., and Strelnikov, V. V. (2010). Amplification of intermethylated sites experimental design and results analysis with AIMS in silico computer software. *Mol. Biol.* 44, 317–325.

Tanemura, A., Terando, A.M., Sim, M., Hoesel, A.Q. Van, Maat, F.G. De, Morton, D.L., and Hoon, D.S.B. (2009). CpG Island Methylator Phenotype Predicts Progression of Malignant Melanoma. *Clin Cancer Res* 15, 1801–1807.

Tellez, C.S., Shen, L., Estécio, M.R.H., Jelinek, J., Gershenwald, J.E., and Issa, J.-P.J. (2009). CpG island methylation profiling in human melanoma cell lines. *Melanoma Res.* 19, 146–155.

Terry, M.B., Ferris, J.S., Pilsner, R., Flom, J.D., Tehranifar, P., Santella, R.M., Gamble, M. V, and Susser, E. (2008). Genomic DNA methylation among women in a multiethnic New York City birth cohort. *Cancer Epidemiol. Biomarkers Prev.* 17, 2306–2310.

Terry, M.B., Delgado-Cruzata, L., Vin-Raviv, N., Wu, H.C., and Santella, R.M. (2011). DNA methylation in white blood cells: Association with risk factors in epidemiologic studies. *Epigenetics* 6, 828–837.

Teschendorff, A.E., Menon, U., Gentry-Maharaj, A., Ramus, S.J., Gayther, S. a, Apostolidou, S., Jones, A., Lechner, M., Beck, S., Jacobs, I.J., et al. (2009). An epigenetic signature in peripheral blood predicts active ovarian cancer. *PLoS One* 4, e8274.

Teschendorff, A.E., West, J., and Beck, S. (2013). Age-associated epigenetic drift: implications, and a case of epigenetic thrift? *Hum. Mol. Genet.* 22, R7–R15.

The Cancer Genoma Atlas (2013). TCGA. Natl. Cancer Inst. Natl. Hum. Genome Res. Inst.

Tobi, E.W., Lumey, L.H., Talens, R.P., Kremer, D., Putter, H., Stein, A.D., Slagboom, P.E., and Heijmans, B.T. (2009). DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum. Mol. Genet.* 18, 4046–4053.

Van Doorn, R., Gruis, N.A., Willemze, R., van der Velden, P.A., and Tensen, C.P. (2005). Aberrant DNA methylation in cutaneous malignancies. *Semin. Oncol.* 32, 479–487.

Veeck, J., and Esteller, M. (2010). Breast cancer epigenetics: from DNA methylation to microRNAs. J. Mammary Gland Biol. Neoplasia 15, 5–17.

Vereecken, P., Cornelis, F., Van Baren, N., Vandersleyen, V., and Baurain, J.F. (2012). A synopsis of serum biomarkers in cutaneous melanoma patients. *Dermatol. Res. Pract.* 12, 1–7.

Virani, S., Colacino, J.A., Kim, J.H., and Rozek, L.S. (2012). Cancer epigenetics: a brief review. *ILAR J*. 53, 359–369.

Waddington, C.H. (1959). Canalization of development and genetic assimilation of acquired characters. *Nature* 183, 1654–1655.

Walter, M.J., Ding, L., Shen, D., Shao, J., Grillot, M., McLellan, M., Fulton, R., Schmidt, H., Kalicki-Veizer, J., O'Laughlin, M., et al. (2011). Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. *Leukemia* 25, 1153–1158.

Wang, S., Dorsey, T.H., Terunuma, A., Kittles, R.A., Ambs, S., and Kwabi-Addo, B. (2012). Relationship between tumor DNA methylation status and patient characteristics in African-American and European-American women with breast cancer. *PLoS One* 7, e37928.

Wang, X.S., Diener, K., Tan, T.H., and Yao, Z. (1998). MAPKKK6, a novel mitogenactivated protein kinase kinase kinase, that associates with MAPKKK5. *Biochem. Biophys. Res. Commun.* 253, 33–37.

Waterland, R. A, and Michels, K.B. (2007). Epigenetic epidemiology of the developmental origins hypothesis. *Annu. Rev. Nutr.* 27, 363–388.

Waterland, R.A., and Jirtle, R.L. (2003). Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell. Biol.* 23, 5293–5300.

Waterland, R.A., Lin, J.-R., Smith, C.A., and Jirtle, R.L. (2006). Post-weaning diet affects genomic imprinting at the insulin-like growth factor 2 (Igf2) locus. *Hum. Mol. Genet.* 15, 705–716.

Welsh, J., and McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* 18, 7213–7218.

Widschwendter, M., Apostolidou, S., Raum, E., Rothenbacher, D., Fiegl, H., Menon, U., Stegmaier, C., Jacobs, I.J., and Brenner, H. (2008). Epigenotyping in peripheral blood cell DNA and breast cancer risk: a proof of principle study. *PLoS One* 3, e2656.

Wild, L., and Flanagan, J.M. (2010). Genome-wide hypomethylation in cancer may be a passive consequence of transformation. *Biochim. Biophys. Acta* 1806, 50–57.

Wilhelm, C.S., Kelsey, K.T., Butler, R., Plaza, S., Gagne, L., Zens, M.S., Andrew, A.S., Morris, S., Nelson, H.H., Schned, A.R., et al. (2010). Implications of LINE1 methylation for bladder cancer risk in women. *Clin. Cancer Res.* 16, 1682–1689.

Williams, P.F., Olsen, C.M., Hayward, N.K., and Whiteman, D.C. (2011). Melanocortin 1 receptor and risk of cutaneous melanoma: a meta-analysis and estimates of population burden. *Int. J. Cancer* 129, 1730–1740.

Winkler, C.A., Nelson, G.W., and Smith, M.W. (2010). Admixture mapping comes of age. *Annu. Rev. Genomics Hum. Genet.* 11, 65–89.

Winnefeld, M., Brueckner, B., Grönniger, E., Stäb, F., Wenck, H., and Lyko, F. (2012). Stable ethnic variations in DNA methylation patterns of human skin. *J. Invest. Dermatol.* 132, 466–468.

Wojdacz, T.K., Thestrup, B.B., Cold, S., Overgaard, J., and Hansen, L.L. (2011). No difference in the frequency of locus-specific methylation in the peripheral blood DNA of women diagnosed with breast cancer and age-matched controls. *Futur. Oncol.* 7, 1451–1455.

Woo, H.D., and Kim, J. (2012). Global DNA hypomethylation in peripheral blood leukocytes as a biomarker for cancer risk: a meta-analysis. *PLoS One* 7, e34615.

Woodson, K., Mason, J., Choi, S., Hartman, T., Tangrea, J., Virtamo, J., Taylor, P.R., and Albanes, D. (2001). Hypomethylation of p53 in peripheral blood DNA Is associated with the development of lung cancer. *Cancer Epidemiol. Biomarkers Prev.* 10, 69–74.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* 2, 241.

Wu, F., Shirahata, A., Sakuraba, K., Kitamura, Y., Goto, T., Saito, M., Ishibashi, K., Kigawa, G., Nemoto, H., Sanada, Y., et al. (2010). Down-regulation of Mus81 as a potential marker for the malignancy of gastric cancer. *Anticancer Res* 30, 5011–5014.

Wu, F., Shirahata, A., Sakuraba, K., Kitamura, Y., Goto, T., Saito, M., Ishibashi, K., Kigawa, G., Nemoto, H., Sanada, Y., et al. (2011). Downregulation of Mus81 as a novel prognostic biomarker for patients with colorectal carcinoma. *Cancer Sci.* 102, 472–477.

Youngson, N. a, and Whitelaw, E. (2008). Transgenerational epigenetic effects. Annu. Rev. Genomics Hum. Genet. 9, 233–257.

Zhang, D., Cheng, L., Badner, J.A., Chen, C., Chen, Q., Luo, W., Craig, D.W., Redman, M., Gershon, E.S., and Liu, C. (2010). Genetic control of individual differences in gene-specific methylation in human brain. *Am. J. Hum. Genet.* 86, 411–419.

Zhang, F.F., Cardarelli, R., Carroll, J., Fulda, K.G., Kaur, M., Gonzalez, K., Vishwanatha, J.K., Santella, R.M., and Morabia, A. (2011). Significant differences in global genomic DNA methylation by gender and race/ethnicity in peripheral blood. *Epigenetics* 6, 623–629.

Zhao, H., Langerød, A., Ji, Y., Nowels, K.W., Nesland, J.M., Tibshirani, R., Bukholm, I.K., Kåresen, R., Botstein, D., Børresen-Dale, A.-L., et al. (2004). Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. *Mol. Biol. Cell* 15, 2523–2536.

Zhao, H., Du, W., Gu, D., Wang, D., Xue, F., Ge, J., Sui, T., and Yang, R. (2009). DNMT3B 579G>T promoter polymorphism and the risk for idiopathic thrombocytopenic purpura in a Chinese population. *Acta Haematol*. 122, 31–35.

Zhu, S., Zhang, H., Tang, Y., Liu, P., and Wang, J. (2012a). DNMT3B polymorphisms and cancer risk: a meta analysis of 24 case-control studies. *Mol. Biol. Rep.* 39, 4429–4437.

Zhu, Z.Z., Sparrow, D., Hou, L., Tarantini, L., Bollati, V., Litonjua, A.A., Zanobetti, A., Vokonas, P., Wright, R.O., Baccarelli, A., et al. (2011). Repetitive element hypomethylation in blood leukocyte DNA and cancer incidence, prevalence, and mortality in elderly individuals: the Normative Aging Study. *Cancer Causes Control* 22, 437–447.

Zhu, Z.Z., Hou, L., Bollati, V., Tarantini, L., Marinelli, B., Cantone, L., Yang, A.S., Vokonas, P., Lissowska, J., Fustinoni, S., et al. (2012b). Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis. *Int. J. Epidemiol.* 41, 126–139.

Zilberman, D., and Henikoff, S. (2007). Genome-wide analysis of DNA methylation patterns. *Development* 134, 3959–3965.



## 8.1 Tablas suplementarias

**Tabla Suplementaria 1.** Líneas celulares comerciales utilizadas del banco de tejidos del Programa de epigenética y biología del cáncer (PEBC) del IDIBELL.

Líneas celulares	Tejido	Тіро
IGR37	piel	melanoma
WM155	piel	melanoma
C32_TG	piel	melanoma
WM793	piel	melanoma
IGR39	piel	melanoma
C32	piel	melanoma
WM_266	piel	melanoma
UACC-257	piel	melanoma
M14	piel	melanoma
MALME-3M	piel	melanoma
UACC-62	piel	melanoma
SK-MEL-28	piel	melanoma
MDA-MB-435	piel	melanoma
SK-MEL-2	piel	melanoma
LOX_IMVI	piel	melanoma
SK-MEL-5	piel	melanoma
ST_BRAF	piel	melanoma
MELST	piel	melanoma
STR	piel	melanoma
HCC-1143	mama	cáncer
HS_578T	mama	cáncer
MDA-MB-231	mama	cáncer
MCF7	mama	cáncer
T-47D	mama	cáncer
BT-549	mama	cáncer
MDA-MB-468	mama	cáncer
MDA-MB-231	mama	cáncer
HCC1143	mama	cáncer
HCC1937	mama	cáncer
UACC-812	mama	cáncer

AIMs	Cromosoma	Alelos	AIMs	Cromosoma	Alelos
rs1934393	1	C/G	rs879780	11	C/T
rs2817611	1	A/G	rs948360	11	A/G
rs3828121	1	C/T	rs28931575	11	ALU/-
rs6684063	1	G/T	rs4034627	12	C/T
rs10498255	2	C/T	rs4076700	12	C/T
rs1470524	2	C/T	rs4762106	12	A/G
rs3860446	2	C/T	rs10492585	13	C/T
rs842634	2	C/T	rs2585901	13	C/T
rs868179	2	A/G	rs10131076	14	A/G
rs1395771	3	A/G	rs1451928	14	G/T
rs1984473	3	C/T	rs9323178	14	A/G
rs6804094	3	A/T	rs10520678	15	C/T
rs9310888	3	A/G	rs9302185	15	C/T
rs1398829	4	A/T	rs1426654	15	A/G
rs9307613	4	A/T	rs1004704	16	A/G
rs10515535	5	A/G	rs10500505	16	A/T
rs257748	5	A/T	rs30125	16	A/G
rs10484578	6	A/G	rs3138523	16	ALU/-
rs6569792	6	A/G	rs4130513	16	C/T
rs6911727	6	C/T	rs10491097	17	A/G
rs9320808	6	A/G	rs2253624	17	G/T
rs10214949	7	A/G	rs1013459	18	A/G
rs10486576	7	C/T	rs12953952	18	A/G
rs4733652	8	C/T	rs798887	19	A/G
rs9325872	8	A/G	rs888861	19	A/G
rs10491654	9	C/T	rs2208139	20	C/T
rs4013967	9	C/T	rs708915	20	A/T
rs10508349	10	A/G	rs2829454	21	A/G
rs1397618	10	A/T	rs138022	22	A/G
rs10501474	11	C/T			

**Tabla Suplementaria 2**. Lista de marcadores informativos de ancestralidad (AIMs) utilizados para determinar la ancestría genética individual, con la posición cromosómica y los alelos.

	Metilación global (% )				
Genotipo	Controles	Pacientes			
CC	3,556	2,437			
CG	2,496	2,289			
GG	2,705	2,250			

**Tabla Suplementaria 3.** Modelo predictivo de regresión de valores promedio de metilación global del ADN de acuerdo a los genotipos del rs545500 en gen *MUS81*.

**Tabla Suplementaria 4.** Características epidemiológicas y tumorales de las pacientes con cáncer de mama analizadas en la plataforma 450K HumanMethylation BeadChip.

ID	Edad	Estadío tumoral	Tumor	Tipo Histológico	RE	RP	Her2	IMC	Fumadora (sí/NO)	Período fumadora
KP031	47	IIA	infiltrante	Ductal	NEG	POS	POS	> 30	SI	En el pasado
KP034	49	IIIA	infiltrante	Ductal	POS	POS	POS	< 25	NO	Nunca
KP043	49	IIB	infiltrante	Papilar	POS	POS	NEG	> 30	NO	Nunca
KP047	45	IIA	infiltrante	Ductal	NEG	NEG	NEG	> 30	NO	Nunca
KP073	68	IIA	infiltrante	Ductal	POS	NEG	POS	> 30	NO	Nunca
KP076	59	I.	infiltrante	Ductal	POS	POS	NEG	> 30	NO	Nunca
KP098	62	IIA	infiltrante	Ductal				> 30	NO	Nunca
KP103	58	I.	infiltrante	Ductal	POS	POS	NEG	< 25	NO	Nunca
KP104	48	IIA	infiltrante	Ductal	POS	POS	NEG	< 25	SI	Actual
KP111	70								NO	Nunca
KP113	77	I.	infiltrante	Ductal	POS	POS	NEG	25 - 29,9	NO	Nunca
KP120	52	IIA	infiltrante	Lobulillar	POS	NEG	NEG	25 - 29,9	SI	Actual
KP129	73								SI	En el pasado
KP132	54				POS	POS	POS	25 - 29,9	NO	Nunca
KP138	52	I.	infiltrante	Ductal	POS	POS	POS	> 30	SI	En el pasado
KP155	58	IIA	infiltrante	Ductal					SI	En el pasado
KP173	52	IIA	infiltrante	Ductal	NEG	NEG	NEG	25 - 29,9	NO	Nunca
KP182	67	IIIA	infiltrante	Lobulillar	POS	POS	NEG	25 - 29,9	NO	Nunca
KP191	69	IIIA	infiltrante	Ductal	POS	POS	NEG	< 25	SI	En el pasado
KP202	63	IIA	infiltrante	Ductal	POS	POS	NEG	> 30	NO	Nunca
KP210	63	I.	infiltrante	Ductal	NEG	NEG	POS	< 25	NO	Nunca
KP214	63	IIA	infiltrante	Ductal	POS	NEG	POS	> 30	NO	Nunca

KP: pacientes con cáncer de mama, KC: controles no afectados, RE: receptor de estrógeno, RP: receptor de progesterona, IMC: índice de masa corporal. Los espacios en blanco significan que no se tiene el dato.

Tejidos pareados	Subtipo	Grado	Estadío	Familiar
N1/T1	TNP	2	llb	No
N2/T2	Luminal A	1	lla	No
N3/T3	Luminal A	1	I	No
N4/T4	Basal	2	lla	No
N5/T5	Luminal B	3	Illa	BRCA2 mut
N6/T6	TNP	3	lla	BRCA2 mut
N7/T7	Basal	3	NA	No
N8/T8	Basal	NA	NA	No
N9/T9	Basal	NA	I	No
N10/T10	Basal	3	I	No
N11/T11	Luminal A	NA	llb	No
N12/T12	Basal	3	lla	No
N13/T13	Basal	2	lla	BRCA2 mut
N14/T14	Basal	3	IIIb	No

**Tabla Suplementaria 5.** Características de los tumores correspondientes a los pares de tejidos mamarios (normal/tumoral) de pacientes con cáncer de mama. TNP: Fenotipo triple negativo, NA: sin datos.

Tabla Suplementaria 6. Comparación de la distribución genómica de los CpGDM con respecto a la representación de las CpG en la plataforma 450K HumanMethylation BeadChip. TSS1500: región a 1500 pb del sitio de inicio de la transcripción; TSS200 región a 200 pb del sitio de inicio de la transcripción. UTR: región no traducida

Región genómica	% CpG representadas en el microarreglo	CpGDM (n)	CpG esperadas (n)	Fisher p-valor
TSS1500	14,97	11	11,5	0,987
TSS200	11,13	8	8,6	0,988
5´UTR	11,41	8	8,8	0,983
1er exón	6,83	1	5,2	0,923
Cuerpo génico	31,61	30	24,3	0,688
3´UTR	3,44	1	2,6	0,971
Intergénica	20,59	18	15,8	0,936

Tabla Suplementaria 7. Comparación de los perfiles de metilación de los 77 CpGDMs en los distintos tipos de muestras analizadas. Se indica el % de CpGDM desmetilados (media  $\beta$ <0,3), hemimetilados (media  $\beta$  entre 0,3 y 0,7) y metilados (media  $\beta$ > 0,7).

	% de CpGDM por rango de metilación					
	< 0.3	0.3 - 0.7	> 0.7			
Sangre						
Pacientes	6.5	57.1	36.4			
Controles	5.2	55.8	39			
Totales	5.2	58.4	36.4			
Tejido mama						
Tejidos sanos	57.1	29.9	13			
Tumores primarios	49.3	36.4	14.3			

% do CoCDM po - do n otilació **Tabla Suplementaria 8.** Comparación de los valores de metilación de los 77 CpGDM entre muestras de ADN de tumores primarios de mama y tejidos mamarios sanos. CHR: cromosoma.

0.0014		UCSC	Media Tejido	Media Tejido	Delta media	Wilcoxon FDR
Срадм	CHR	REFGENE	tumoral	sano	(tumoral-sano)	p-valor
cg01015663	1	TCEA3	0,4470	0,3294	0,1176	0,0010
cg01229567	1	MIB2	0,2531	0,2694	-0,0164	0,2430
cg04400047	1	UBIAD1	0,1751	0,0993	0,0757	2.56e-05
cg06432479	1	TAL1	0,1328	0,1491	-0,0163	0,1130
cg10159215	1	LPPR5	0,2023	0,1574	0,0449	0,1020
cg14024502	1	MAP3K6	0,5722	0,4881	0,0841	0,0009
cg15452381	1	HIVEP3	0,7397	0,7552	-0,0155	0,4640
cg16727538	1	C1orf213	0,0591	0,0434	0,0157	0,0028
cg19246761	1	MIB2	0,1675	0,1825	-0,0151	0,1240
cg26251270	1	GPX7	0,1484	0,0499	0,0985	0,0014
cg26750487	1	CR1L	0,1844	0,1110	0,0733	0,0087
cg02537909	2	COBLL1	0,0857	0,0795	0,0063	0,2949
cg02738156	2		0,5256	0,5564	-0,0307	0,4646
cg09408768	2	KCNJ3	0,2183	0,1561	0,0623	0,0044
cg17165836	2	AFF3	0,3525	0,2267	0,1258	0,6386
cg24130711	2	CCDC85A	0,1422	0,1670	-0,0248	0,0094
cg26175971	2	CYP2/A1	0,1396	0,0927	0,0469	0,0001
cg26874367	2		0,3149	0,4262	-0,1113	0,0014
cg01615258	3	0450	0,1012	0,1674	-0,0662	2.26e-06
cg01814969	3	QARS	0,1267	0,1134	0,0133	0,0117
cg15450445	3	00004	0,7343	0,6261	0,1081	3.31e-06
cg24840062	3	CDCP1	0,5728	0,5495	0,0233	0,3715
cg∠5616514	3		0,4697	0,5654	-0,0956	4.936-05
cgu3002688	4	81.04040	0,7234	0,7741	-0,0507	0,0119
cg18860310	4	SLUTUA6	0,3999	0,4241	-0,0242	0,3231
cg26994377	4		0,3062	0,2642	0,0421	0,2839
000008540	5		0,0270	0,0511	-0,0241	0,1248
cg01204911	5		0,7397	0,7720	-0,0323	0,0602
cg06527989	5		0,1532	0,0983	0,0550	0,0117
cg07475151	5	LUC100200100	0,1200	0,0622	0,0463	0,0046
cg11953913	5	C501132	0,7701	0,7915	-0,0213	0,0137
cg21550107	5		0,5409	0,3721	0,1000	3.310-00
cg02174359	6	IVIRP 5 TOA	0,1073	0,1036	0,0035	0,7942
cg04334016	6	CINKSKS	0,3796	0,2523	0,1274	4.966-05
cg09639771	0	CDV4	0,5994	0,0440	-0,0454	0,0087
cg02377005	7		0,1341	0,0059	0,0002	0,0094
cg03701471	7		0,2752	0,2004	0,0097	0,0717
cg15602580	7		0,0912	0,9222	-0,0310	0,2080
cg18967180	7		0,0300	0,0775	0,0193	0,0037
cg21252523	7	DENNDZA	0,1072	0,1003	-0.0659	0,0025
cg27403098	7	KIA A 1908	0,0000	0,0007	0.0819	0,0005
cg14681767	8	ARHGEE10	0,4342	0,0320	0.0417	0,0005
cg17588094	8	ANNOLITO	0,7775	0,0030	0.0288	0,0020
cq05616472	9	EHMT1	0 1022	0.0883	0.0139	0.0729
cg14223444	9	TSTD2	0.3940	0.3573	0.0367	0,2067
cq01311537	10	C10orf128	0.8361	0.9221	-0.0860	1.38e-05
cq08560387	10	TSPAN14	0.5360	0.5584	-0.0224	0.4646
cg14770293	10		0.1998	0.1228	0.0770	0.0002
cg24168991	10	ITPRIP	0.7206	0.6143	0.1062	0.0002
cq03611487	11	LOC100126784	0.4983	0.4471	0.0512	0.1027
cq10141801	11	GUCY2E	0,0942	0,0774	0,0168	0,0008
cq13100962	11		0,3106	0,1414	0,1692	7.85e-06
cg17679104	11	STK33	0,2364	0,1590	0,0775	0,6457
cg22623080	11	AMOTL1	0,1180	0,0865	0,0315	0,0090
cg23460961	11		0,4496	0,5994	-0,1498	0,0001
cg04890607	12	HMGA2	0,6053	0,6526	-0,0472	0,3779
cg27292547	12		0,6016	0,6440	-0,0424	0,7645
cg11398020	13	KLF5	0,0799	0,0509	0,0290	0,0061
cg01972418	14	PAX9	0,3000	0,2011	0,0989	0,0837
cg18581173	15	CT62	0,2023	0,2606	-0,0583	3.45e-05
cg20172862	15		0,0598	0,0572	0,0026	0,7645
cg24359188	15	BUB1B	0,1917	0,1489	0,0427	0,0004
cg26568226	15	CYFIP1	0,8389	0,8959	-0,0570	0,0201
cg04470044	16	WFDC1	0,5511	0,5362	0,0149	0,3714
cg10155261	16		0,1138	0,1161	-0,0023	0,4646
cg26591162	16	SRL	0,7486	0,7025	0,0462	0,0094
cg07777703	17	TUBG2	0,1734	0,1219	0,0515	7.85e-06
cg07848706	17		0,1061	0,1171	-0,0110	0,5149
cg08960549	17		0,5824	0,7393	-0,1569	5.29e-05
cg09580608	17	GNA13	0,0750	0,0777	-0,0027	0,8761
cg22163463	17	PITPNM3	0,3987	0,4058	-0,0071	0,6528
cg01823541	19	GNG7	0,1044	0,0886	0,0157	0,0117
cg03363633	19	TYROBP	0,4796	0,4886	-0,0090	0,7645
cg22313519	19	KIAA1683	0,3069	0,1439	0,1630	3.31e-06
cg20477147	20		0,2508	0,2258	0,0250	0,5078
cg26468205	20	PCMTD2	0,0347	0,0327	0,0019	0,7878

**Tabla Suplementaria 9.** Características asociadas a los criterios de selección establecidos de los CpGDM candidatos. TP: tumor primario, TNM: tumor normal de mama. Las X indican cumplimiento del criterio. Las flechas indican disminución del nivel evaluado. \*algunas publicaciones lo asocian a progresión en melanoma cutáneo.

	CpGDM candidatas				
	cg01229567	cg19246761	cg14024502	cg26568226	
Ubicación cromosómica	1p36.33	1p36.33	1p36.11	15q11.2	
UCSC REFGENE	MIB2	MIB2	MAP3K6	CYFIP1	
Contexto génico	promotor	promotor	promotor	5´UTR	
Gen asociados a cáncer	melanoma*	melanoma*	mama	mama	
∆beta >10%	Х	Х			
CpG hipermetilada				Х	
Expresión diferencial en PBMC (cáncer de mama vs sanos)	X↓	×↓	X↓		
Metilación diferencial (TP vs TNM)			Х	Х	
Expresión diferencial (TP vs TNM)	X↓	X↓	X↓		

## 8.2 Figuras suplementarias

**Figura suplementaria 1.** Agrupamiento jerárquico basado en los valores de metilación de los 77 CpGDM de muestras de tumores primarios mamarios (TP) y tejidos mamarios sanos (TMN). Se utilizó el método de aglomeración "completo" usando distancias "euclidianas". Los CpGDM fueron ordenados por la diferencia de las medias de los valores betas entre tejidos tumorales y tejidos sanos. El nivel de metilación de los CpGDM está codificado en colores (verde: menor nivel de metilación; rojo: mayor nivel de metilación). Los círculos azules identifican muestras de tejido mamario sano dentro del cluster de tejidos tumoral.



**Figura suplementaria 2.** Secuenciación con bisulfito de los clones individuales a las regiones flanqueantes al sitio CpG evaluado por la sonda cg14024502 en el gen *MAP3K6*. Cada cuadrado corresponde a un sitio CpG en el fragmento amplificado. KC: controles; KP: pacientes. Cuadrados oscuros: metilados; cuadrados blancos: no metilados. Se indica el promedio de metilación de la región evaluada en cada individuo.



Región TSS1500 MAP3K6

**Figura suplementaria 3.** Secuenciación con bisulfito de los clones individuales a las regiones flanqueantes a los sitios CpGs evaluados por las sondas cg19246761 y cg01229567 en el gen *MIB2*. Cada cuadrado corresponde a un sitio CpG en el fragmento amplificado. KC: controles; KP: pacientes. Cuadrados oscuros: metilados; cuadrados blancos: no metilados. Se indica el promedio de metilación de la región evaluada en cada individuo.



MIB2