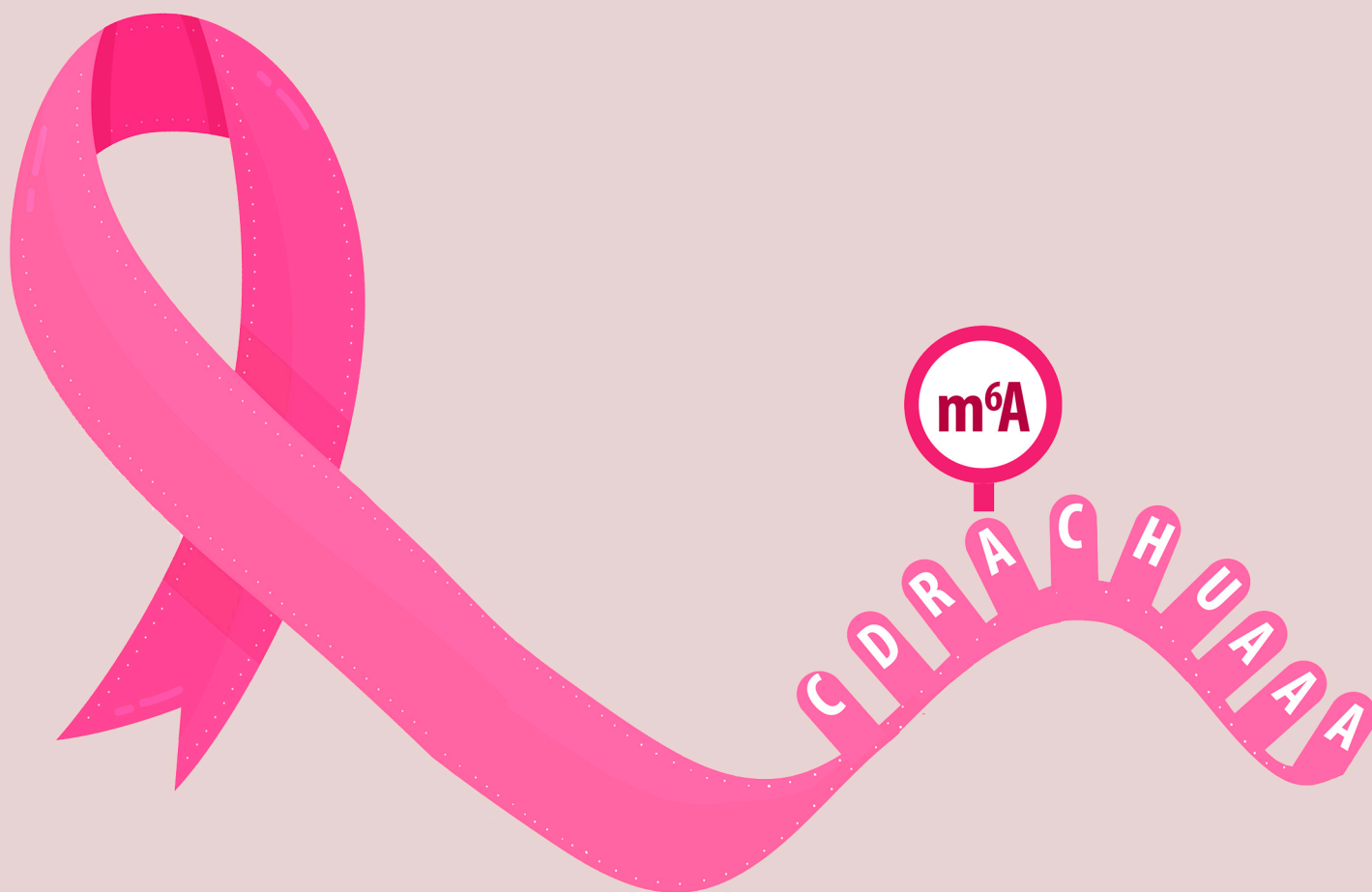


Descubrimiento de nuevas dianas terapéuticas en cáncer de mama basado en el análisis de metilación del ARNm

TESIS DOCTORAL 2022
TAMARA SOLEDAD KLEINBIELEN



eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

eman ta zabal zazu



Universidad del País Vasco Euskal Herriko Unibertsitatea

Departamento de Bioquímica y Biología Molecular

Grupo de Investigación BIOMICs

Descubrimiento de nuevas dianas terapéuticas en cáncer de mama basado en el análisis de metilación del ARNm

Memoria presentada por

Tamara Soledad Kleinbielen

Para optar al grado de

Doctora por la Universidad del País Vasco/ Euskal Herriko Unibertsitatea

Trabajo dirigido por

Dr. Felix Olasagasti

Dra. Elena Beristain Mendizabal

Vitoria-Gasteiz, 2022

Este trabajo ha sido llevado a cabo gracias a la financiación por parte del Departamento de Educación, Universidades e Investigación del Gobierno Vasco de los grupos de investigación BIOMICs y Microfluidics & BIOMICs cluster (IT- 833-13 e IT-1271-19, respectivamente), y a la colaboración con el grupo liderado por África García Orad de la Facultad de Medicina de la Universidad del País Vasco UPV/EHU.

De igual forma, este trabajo ha sido posible gracias a la concesión, por parte del Departamento de Educación del Gobierno Vasco, de una beca predoctoral para la realización de esta Tesis Doctoral, la cual he disfrutado entre los años 2.018 y 2.022.

AGRADECIMIENTOS

En estas líneas me gustaría plasmar mis agradecimientos a todas las personas que me han ayudado, animado y acompañado en todo el proceso de la tesis.

A Marian Martínez de Pancorbo, por confiar en mí e invitarme a realizar la tesis en BIOMICS. Gracias por haber invertido tiempo y dedicación en mí y en este proyecto. Gracias por todo lo que me has enseñado y por todos los consejos dados a lo largo de todos estos años.

A mis directores. Gracias Felix Olasagasti por la guía, consejos y aliento para llevar a cabo esta tesis. Si hoy puedo estar escribiendo este apartado es, en parte, gracias a ti. Gracias Elena Beristain por estar y porque sé que puedo contar contigo.

A Ana Dopazo, Ana Quintas y Quique Vázquez del CNIC. Si esta tesis ha podido llegar a buen puerto ha sido, en gran parte, gracias a vuestra ayuda. Gracias por vuestros consejos, por ayudarnos con la optimización de la técnica, por secuenciar las muestras y realizar todo el análisis informático. Gracias por vuestro tiempo, vuestros mensajes y videollamadas. Gracias por haber trabajado a contrareloj para obtener los resultados.

A África García Orad por haber apoyado este proyecto desde el comienzo y por aportarnos todos los recursos necesarios para iniciarlo.

A Isabel Guerra y a Amparo Viguri, gracias por tomarse el tiempo revisar los datos asociados a las muestras y por haberme respondido las dudas sobre la anatomía patológica del tejido mamario.

A los que fueron y son mis compañer@s de laboratorio: Miriam, Maite, Leire, Eva, Cate, Belén, Dani, Patri, Andrés, David, y en especial a Endika, con quien iniciamos este proyecto y nos sumergimos en el mundo de la epitranscriptómica. Gracias a tod@s por los buenos momentos, las charlas, por animarme, por escucharme y por intentar buscar respuestas a mis preguntas.

A mi familia. Gracias a mis padres por apoyarme y animarme en este sueño de venir a estudiar a España. Gracias porque, aunque estamos lejos, vosotros siempre estuvisteis y estáis muy cerca. Gracias a mi madre por enseñarme a cómo enfrentar las circunstancias adversas de la vida y a mis hermanos por estar siempre.

A mis familiares y amigos de Argentina. Gracias por acompañarme en este proceso desde la distancia, gracias por cada mensaje y llamada.

AGRADECIMIENTOS

A Egoitz, gracias por llegar a mi vida en el momento indicado. Gracias por escucharme, aconsejarme, animarme y revisar la bibliografía. Gracias por ayudarme a llegar a la meta y gracias porque podemos correr juntos esta carrera de la vida. Gracias a Gerardo, Montse, Naroa y Markel. Gracias por recibirme en vuestra casa con los brazos abiertos y por hacerme una más de vuestra familia. Gracias Naroa por el diseño de la portada, ¡me ha encantado!

A la gran familia que tengo en Vitoria-Gasteiz, vuestro apoyo ha sido muy importante para mí.
A mi querido amigo Jesús.

¡Muchas gracias!

Tamara S. Kleinbielen

RESUMEN

N⁶-metiladenosina (m⁶A) es una de las modificaciones más abundante del ARNm en los organismos eucariotas. Distintos estudios han revelado que la modificación m⁶A está involucrada en la regulación de la mayoría de los procesos del metabolismo del ARN. Sin embargo, aún no existe mucho conocimiento acerca del impacto que esta modificación tiene sobre la regulación de la expresión génica, su función en los procesos celulares y el efecto que puede ejercer en el desarrollo y progresión de ciertas enfermedades, como, por ejemplo, el cáncer de mama. En este estudio, exploramos la presencia de la modificación m⁶A en el cáncer de mama, específicamente en los subtipos moleculares luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo, con el fin de determinar el estado de la metilación y de su posible relación con futuras dianas terapéuticas.

Para ello, en primer lugar, hemos determinado el porcentaje de la metilación m⁶A global presente en las muestras de tejido de pacientes con cáncer de mama y evaluado la expresión génica de las metiltransferasas y desmetilasas. Los resultados de este estudio revelan que el porcentaje de metilación m⁶A global es significativamente menor en las muestras de tejido tumoral en comparación con sus correspondientes tejidos sanos. El análisis de la expresión de los principales genes modificadores de m⁶A analizados, es decir, *METTL3*, *METTL14*, *WTAP*, *FTO* y *ALKBH5*, ha mostrado que estos genes presentan una expresión reducida en muestras tumorales con respecto a sus correspondientes tejidos sanos. Además, el perfil de expresión génica de metiltransferasa y desmetilasas es característico de cada subtipo molecular de cáncer de mama aquí estudiado, así como las correlaciones entre dichos perfiles, lo que podría ser útil para la clasificación molecular de estos tumores.

Por otro lado, se realizó una búsqueda *in silico* para identificar los m⁶A-SNPs asociados al cáncer de mama y evaluar sus posibles efectos sobre la expresión génica. Se identificaron 981 m⁶A-SNPs después de comparar 69.713 SNPs relacionados con el cáncer de mama con 301.875 m⁶A-SNPs. De estos, solo cuatro m⁶A-SNPs presentaron un efecto *cis*-eQTL: rs1801270 en *CDKN1A*, rs76563149 en *ZNF354A*, rs9379084 en *REEB1* y rs11614913 en *MIR196A2*, y todos ellos, excepto rs9379084, mostraron una expresión génica alterada en el tejido mamario. Para validar los resultados obtenidos, se evaluó *in vitro* la expresión de los genes donde se encontraban los tres m⁶A-SNPs que presentaron diferencias significativas en el análisis *in silico*, utilizando muestras de tejido de pacientes con cáncer de mama. Nuestro análisis mostró que solo el m⁶A-

SNP rs76563149, ubicado en el gen *ZNF354A*, presenta diferencias en las frecuencias alélicas y una baja expresión génica en las muestras tumorales en comparación las muestras de tejido sano.

Finalmente, mediante el uso de la técnica MeRIP-seq, se ha determinado el perfil de metilación m⁶A presente en todo el transcriptoma de cada uno de los cinco subtipos moleculares de cáncer de mama estudiados. Ello ha permitido poner de manifiesto 1.703 picos m⁶A metilados y determinar tanto los picos m⁶A específicos de cada subtipo molecular como los compartidos, el número de picos por transcriptos, la distribución de los picos m⁶A en todo el transcriptoma y en los cromosomas. Además, hemos realizado un análisis de ontología genética (GO) y de las rutas metabólicas de KEGG para revelar las funciones potenciales de los primeros 20 genes que presentaban una puntuación de detección más elevada en cada subtipo molecular y hemos evaluado la expresión génica *in silico* de los primeros cinco genes.

Hasta donde sabemos, este es el primer estudio que describe el perfil de metilación m⁶A presente en los distintos subtipos moleculares del cáncer de mama. Los resultados obtenidos muestran la presencia de metilación m⁶A en numerosos ARNms de genes altamente relacionados con el inicio y desarrollo tumorales lo que genera nuevas vías de estudio en esta enfermedad.

ÍNDICE

AGRADECIMIENTOS	15
RESUMEN	17
ÍNDICE	19
1. INTRODUCCIÓN	23
1.1. Cáncer de mama	25
1.1.1. Estructura de la mama	25
1.1.2. Cáncer de mama	27
1.1.2.1. Clasificación histológica	28
1.1.2.2. Grado histológico	28
1.1.2.3. Estratificación tumoral TNM.....	29
1.1.2.4. Invasión de vasos sanguíneos y linfáticos.....	29
1.1.2.5. Multicentricidad del tumor y necrosis tumoral	29
1.1.2.6. Clasificación subrogada.....	30
1.1.2.7. Clasificación molecular.....	30
1.2. Modificaciones en el ARN	32
1.2.1. Modificación m ⁶ A.....	32
1.2.2. Enzimas involucradas en el proceso de metilación/desmetilación	34
1.2.2.1. Proteínas metiltransferasas o “ <i>writers</i> ”	34
1.2.2.2. Proteínas desmetilasas o “ <i>erasers</i> ”	35
1.2.2.3. Proteínas de unión o “ <i>readers</i> ”	36
1.2.3. Funciones biológicas de la modificación m ⁶ A	37
1.2.4. Métodos de análisis del m ⁶ A.....	37
1.3. La metilación m⁶A y el cáncer de mama	42
1.3.1. Metiltransferasas, desmetilasas y proteínas de unión al m ⁶ A en el cáncer de mama	43
1.3.2. m ⁶ A asociados a SNP (m ⁶ A-SNP)	45
1.3.3. Perfil de metilación m ⁶ A en todo el transcriptoma.....	46
2. HIPÓTESIS Y OBJETIVOS	47
2.1. Hipótesis	49
2.2. Objetivos	50

3. MATERIALES Y MÉTODOS	51
3.1. Muestras biológicas	53
3.2. Extracción y cuantificación de ARN total	53
3.2.1. Extracción de ARN total.....	53
3.2.2. Cuantificación y análisis de calidad del ARN total.....	56
3.3. Análisis de la metilación m⁶A	57
3.3.1. Determinación del porcentaje de metilación m ⁶ A global	57
3.3.2. Análisis de expresión génica de las metiltransferasas y desmetilasas.....	59
3.3.2.1. Generación del ADNc a partir del ARN total.....	59
3.3.2.2. Cuantificación del ADNc	59
3.3.2.3. Diseño de los cebadores	60
3.3.2.4. Análisis de eficiencia.....	61
3.3.2.5. RT-qPCR.....	62
3.3.2.6. Análisis de expresión génica.....	62
3.3.2.7. Análisis estadístico	63
3.4. Identificación <i>in silico</i> y análisis de expresión <i>in vitro</i> de m⁶A-SNPs	63
3.4.1. Determinación de m ⁶ A-SNPs relacionados con el cáncer de mama	63
3.4.2. Análisis de loci de rasgos cuantitativos de expresión de los m ⁶ A-SNPs	64
3.4.3. Análisis de expresión génica <i>in silico</i>	64
3.4.4. Validación de las posiciones m ⁶ A-SNPs.....	65
3.4.4.1. Muestras a analizar y diseño de las parejas de cebadores.....	65
3.4.4.2. Secuenciación Sanger	66
3.4.4.3. Análisis de expresión génica.....	67
3.4.4.4. Análisis estadísticos.....	68
3.5. Determinación de posiciones m⁶A metiladas mediante la técnica MeRIP	68
3.5.1. Fragmentación del ARN total.....	68
3.5.2. Inmunoprecipitación	69
3.5.3. Cuantificación de las muestras.....	71
3.5.4. Generación de la librería	71
3.5.5. Cuantificación de la librería y evaluación de la distribución de tamaños.....	72
3.5.6. Análisis de secuenciación masiva	72
3.5.7. Análisis de ontología génica y de rutas metabólicas	73
3.5.8. Análisis de expresión génica <i>in silico</i>	73

4. RESULTADOS	75
4.1. Muestras analizadas en el estudio	77
4.2. Análisis de la metilación m⁶A	78
4.2.1. Determinación del porcentaje de metilación m ⁶ A global	78
4.2.2. Análisis de expresión génica.....	81
4.2.3. Análisis de correlación	90
4.3. Identificación <i>in silico</i> y análisis de expresión <i>in vitro</i> de m⁶A-SNPs relacionados con el cáncer de mama	91
4.3.1. m ⁶ A-SNPs relacionados al cáncer de mama.....	91
4.3.2. Análisis <i>cis</i> -eQTL	92
4.3.3. Análisis <i>in silico</i> de la expresión génica	95
4.3.4. Validación de los m ⁶ A-SNPs en muestras de cáncer de mama.....	97
4.3.5. Análisis de expresión génica para las posiciones m ⁶ A-SNPs	98
4.4. Determinación de posiciones m⁶A metiladas	100
4.4.1. Posiciones m ⁶ A metiladas en el cáncer de mama.....	100
4.4.2. Análisis de ontología génica (GO) y de las rutas metabólicas de KEGG.....	105
4.4.3. Análisis <i>in silico</i> de expresión génica.....	108
5. DISCUSIÓN	113
5.1. Análisis de la metilación m ⁶ A.....	115
5.2. Identificación <i>in silico</i> y análisis de expresión <i>in vitro</i> de m ⁶ A-SNPs relacionados con el cáncer de mama	118
5.3. Determinación de posiciones m ⁶ A metiladas.....	121
6. CONCLUSIONES	127
7. BIBLIOGRAFÍA	131

1. INTRODUCCIÓN

1.1. CÁNCER DE MAMA

1.1.1. Estructura de la mama

La mama es un órgano glandular especializado en la secreción de leche y representa una característica sexual secundaria del sexo femenino (1). Es un órgano constituido por piel, tejido adiposo y conectivo subcutáneo que envuelve a la glándula mamaria. Cada glándula mamaria está formada por 15 a 20 lóbulos que se dividen a su vez en lobulillos más pequeños en donde se encuentran los alveolos, o acinos, formados por células epiteliales capaces de secretar leche (Figura 1.1.1A). Estas estructuras están conectadas por tubos denominados ductos o conductos galactóforos, que irán confluyendo en otros de mayor tamaño hasta desembocar en el pezón. El pezón está rodeado por un área de piel pigmentada denominada areola (1–3).

Todo el tejido mamario está vascularizado principalmente por tres sistemas: las ramas perforantes de la arteria torácica interna que irriga la porción medial de la mama, la rama de la arteria axilar que irriga la porción lateral y la rama de las arterias intercostales que contribuyen a la irrigación de toda la mama (1). Las venas siguen el mismo camino que las arterias mencionadas para drenar en las venas axilares, torácica interna e intercostales. Por otra parte, el sistema linfático de la mama parte de la piel, la glándula y los conductos galactóforos y se dirige hacia la areola formando una red subareolar. Desde este plexo subareolar emergen los linfáticos eferentes, que van a drenar a los ganglios linfáticos situados principalmente en la axila, aunque también en los ganglios paraesternales y supraclaviculares (1). Este drenaje linfático tiene suma importancia, especialmente en el aspecto patológico debido a que los carcinomas de mama tienen a expandirse a través de los vasos linfáticos.

Las mamas son órganos estructuralmente dinámicos, que varían según la edad, el ciclo menstrual y el estado reproductivo de la mujer. Comienzan a formarse durante el desarrollo fetal, con un engrosamiento en el área del pecho que se denomina cresta mamaria o líneas lácteas. Cuando un bebé nace, ya se ha formado el pezón y la parte inicial del conducto galactóforo. Al empezar la pubertad, comienzan a desarrollarse las mamas de las mujeres por acción de las hormonas estrogénicas y progesterona (2). La glándula mamaria aumenta de tamaño, sobre todo, por una mayor deposición de tejido adiposo en el interior de la glándula. También, los conductos se extienden y se ramifican en conductos secundarios, en cuyos terminales se formarán los lobulillos. Este desarrollo continúa gradualmente hasta la edad adulta (aproximadamente los 35 años). Aunque con cada ciclo menstrual se produce una

pequeña remodelación de la mama, no es hasta el embarazo y la lactancia cuando se produce una remodelación completa, transformándose en un órgano maduro y completamente funcional. Esto ocurre en respuesta a las hormonas, estrógenos, progesterona y prolactina, que producen la síntesis *de novo* de nuevos conductos, el alargamiento de los conductos existentes, la morfogénesis alveolar, la diferenciación secretora, y el aumento de la vascularización (4). Una vez producido el parto, los niveles de progesterona caen abruptamente y, por otro lado, aumentan los de prolactina, insulina, corticoides y hormona de crecimiento, permitiendo la

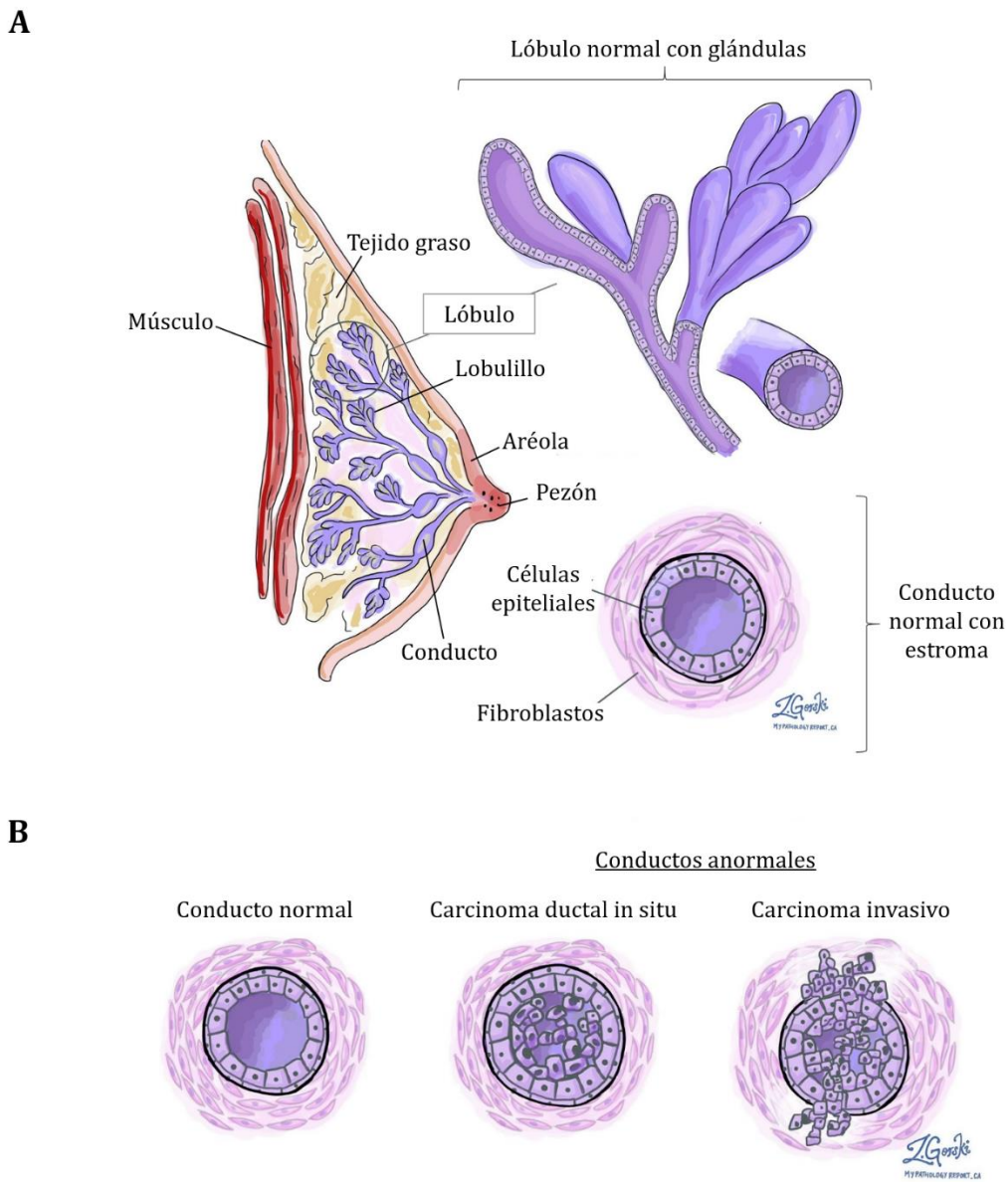


Figura 1.1.1. A- Anatomía de la mama femenina. B- Anatomía de un conducto normal y de conductos anormales. Figuras modificadas de *My Pathology Report* (<https://www.mypathologyreport.ca/es/>).

secreción láctea. El mantenimiento de la lactancia se debe a un mecanismo mediado por la oxitocina. Finalmente, en la menopausia con el cambio del entorno hormonal, el componente glandular de la mama involuciona y es reemplazado por tejido conjuntivo y adiposo.

1.1.2. Cáncer de mama

El cáncer de mama es una enfermedad heterogénea, tanto biológica como clínicamente. El cáncer de mama es el cáncer más comúnmente diagnosticado en las mujeres en el todo mundo (5). Su incidencia, mortalidad y tasas de supervivencia varían considerablemente entre diferentes áreas geográficas, lo que podría deberse a muchos factores como la estructura de la población, el estilo de vida, los factores genéticos y el medio ambiente (6).

Según los datos recogidos por el Sistema Europeo de Información del Cáncer, en España se diagnosticaron un total de 34.088 nuevos casos de cáncer de mama en el año 2.020, siendo el tumor más frecuente entre las mujeres. La probabilidad estimada de desarrollar cáncer de mama es 1 de cada 8 mujeres, apareciendo generalmente entre los 35 - 80 años. En la Comunidad Autónoma Vasca, el cáncer de mama presenta una de tasa de incidencia de 130 casos por cada 100.000 habitantes, siendo la franja de los 35-65 años la de mayor incidencia, debido a que en esas edades es cuando se producen los cambios hormonales de los períodos de peri y post menopausia (datos obtenidos de Plan oncológico de Euskadi 2.018-2.023, <https://www.euskadi.eus/inicio/>).

El cáncer de mama se origina en las células epiteliales de los conductos en un 85% de los casos y en los lóbulos en un 15% de los casos. Al comienzo, el tumor se encuentra confinado en el conducto o lóbulo (cáncer de mama *in situ*) (Figura 1.1.1B). Sin embargo, con el paso del tiempo, este tumor *in situ* puede progresar e invadir el tejido mamario circundante (cáncer de mama invasivo) (Figura 1.1.1B), y a continuación propagarse a los ganglios linfáticos cercanos (metástasis regional) u a otros órganos del organismo (metástasis distante). Por tal motivo, la detección temprana y la selección del tratamiento adecuado es de suma importancia para aumentar la supervivencia de las pacientes que padecen esta enfermedad. En este sentido, la selección del tratamiento presenta ciertas dificultades debido a la heterogeneidad biológica, histológica y clínica del cáncer de mama, por lo que, se están llevando a cabo diversos estudios con el fin de poder caracterizarlo aún más e identificar nuevas dianas terapéuticas (7-9). A

continuación, realizaremos una breve descripción de los factores pronósticos y predictivos utilizados en el cáncer de mama.

1.1.2.1. Clasificación histológica

La clasificación histológica del cáncer de mama se basa en una amplia gama de criterios, como por ejemplo tipo de célula tumoral, secreción extracelular, características arquitectónicas, etc. (10). Hay más de 20 tipos histológicos diferentes de cánceres de mama invasivos. El más común es el carcinoma invasivo de ningún tipo especial (IBC-NST, por sus siglas en inglés *Invasive Breast Cancer No Special Type*), anteriormente conocido como carcinoma ductal invasivo, que representa el 70 - 80% de todos los cánceres invasivos. El carcinoma lobulillar invasivo es el segundo más frecuente y comprende aproximadamente el 10% de los casos. El resto son tipos histológicos menos comunes, como carcinoma mucinoso, cribiforme, micropapilar, papilar, tubular, medular, metaplásico, apocrino, neuroendócrino y quístico adenoide. La clasificación histológica del cáncer de mama tiene relevancia pronóstica y en algunos casos, permite reconocer si el tumor es primario, es decir, si el tumor se ha originado en la mama (11).

1.1.2.2. Grado histológico

El sistema de gradación histológica de Bloom-Richardson (Nottingham) para el cáncer de mama es hoy día el sistema más ampliamente utilizado en todo el mundo. Este sistema de gradación histológica se basa en el grado de diferenciación del tejido tumoral (12). Para ello, se tienen en cuenta tres características morfológicas: grado de formación de túbulos o glándulas (parecido del tumor a la estructura celular normal), pleomorfismo nuclear (diferencias entre los núcleos de las células tumorales y las células normales), y recuento mitótico (incremento en la tasa de división celular). Este sistema evalúa cada característica con un sistema de puntuación numérica de 1 a 3, que posteriormente se suman para tener la clasificación Nottingham, que va del I al III. En este sentido, el grado I (bien diferenciado) se corresponde con una puntuación de 3 a 5, el grado II (moderadamente diferenciado) con una puntuación de 6-7 y el grado III (pobrementemente diferenciado) con una puntuación de 8-9. Los cánceres de grado I tienden a ser menos agresivos, mientras que los de grado III tienden a ser más agresivos. El grado histológico es un poderoso factor de pronóstico (10).

1.1.2.3. *Estratificación tumoral TNM*

La estratificación TNM, ampliamente utilizada en el cáncer de mama, utiliza información tanto clínica como patológica del tamaño del tumor (T), el estado de los ganglios linfáticos regionales (N) y la presencia o ausencia de metástasis (M) para determinar la extensión anatómica del cáncer, el pronóstico y el tratamiento más adecuado (13). En este sistema se asigna un valor para T, N y M, que posteriormente se combina para estratificar la enfermedad en una de las etapas generales. Estas etapas van desde I hasta IV, donde la etapa IV es la más alta y significa que el cáncer está más avanzado que en las etapas más bajas.

1.1.2.4. *Invasión de vasos sanguíneos y linfáticos*

La invasión del sistema vascular y del sistema linfático se produce cuando las células del cáncer de mama irrumpen en los vasos sanguíneos y en los vasos linfáticos (14). También se conoce como invasión linfovascular o vascular-linfática. La invasión de las células tumorales al sistema vascular y linfático aumenta el riesgo de que el cáncer se disemine a otras partes del cuerpo.

1.1.2.5. *Multicentricidad del tumor y necrosis tumoral*

La mama puede dividirse en cuadrantes si se trazan dos rectas, una vertical y otra horizontal, que se intersectan perpendicularmente a nivel del pezón. De esta manera, se pueden identificar cuatro cuadrantes: el cuadrante superior externo (CSE), el cuadrante superior interno (CSI), el cuadrante inferior interno (CII) y el cuadrante inferior externo (CIE) (15). Teniendo esto en cuenta, la multicentricidad del tumor se define como la presencia de 2 o más focos tumorales en distintos cuadrantes de la misma mama o a más de 5 centímetros del foco primario (16,17). Esta característica se ha asociado con un peor pronóstico de la enfermedad.

La necrosis es una forma de muerte celular causada por factores externos a la célula, como la hipoxia. Se identifica en biopsias tumorales al observar células del cáncer de mama muertas que, a menudo, se limitan a una pequeña área de la muestra. La presencia de necrosis generalmente se asocia con formas agresivas de cáncer de mama. Por lo que, su identificación puede proporcionar información del pronóstico y se tiene en cuenta para la elección del tratamiento (18).

1.1.2.6. Clasificación subrogada

La clasificación subrogada del cáncer de mama se usa generalmente en la clínica para clasificar a los tumores en subtipos intrínsecos subrogados y se basan en la expresión histológica e inmunohistoquímica de las siguientes proteínas: receptor de estrógeno (ER), receptor de progesterona (PR), receptor del factor de crecimiento epidérmico humano 2 (HER2) y el marcador de proliferación Ki-67 (19). Los ER y PR son receptores hormonales. Son marcadores de pronóstico y predictivos de la respuesta a la terapia hormonal y se expresan aproximadamente en un 75% de todos los cánceres de mama. La mayoría de los cánceres ER+ también son PR+. Sin embargo, un pequeño porcentaje de cánceres de mama muestran positividad para un solo receptor hormonal. En estos casos, los tumores parecen ser más agresivos y menos sensibles a la terapia hormonal en comparación con los cánceres positivos para ER/PR (10). En cuanto al marcador HER2, aproximadamente el 15 % de los cánceres de mama lo sobreexpresan. El estado de HER2 se analiza mediante la combinación de las técnicas de inmunohistoquímica e hibridación del ADN *in situ* (FISH). La sobreexpresión de HER2 se asocia con un curso clínico agresivo y mal pronóstico, pero también predice la respuesta a los tratamientos dirigidos anti-HER2 (Trastuzumab o Herceptin, Pertuzumab) (10). El 10% restante de los cánceres de mama no expresan ninguno de estos tres marcadores y se denominan cánceres de mama triple negativos. Este tipo de cáncer es, en general de alto grado y se asocia con un mal pronóstico. Por otra parte, el marcador de proliferación Ki-67 sirve para evaluar la actividad proliferativa de un tumor. Si los valores de Ki-67 son altos, es probable que el tumor crezca y se propague más rápido.

1.1.2.7. Clasificación molecular

En las últimas décadas, se han concentrado muchos esfuerzos para complementar la clasificación morfológica del carcinoma de mama con parámetros moleculares que pueden proporcionar una apreciación más clara de la heterogeneidad del cáncer de mama y una mejor predicción del comportamiento del tumor para mejorar las estrategias terapéuticas. En este sentido, Perou *et al.* (2000) clasificaron al cáncer de mama molecularmente en distintos subgrupos basándose en similitudes de los perfiles de expresión génica obtenidos mediante la técnica de microarrays (20). Los principales genes responsables de esta clasificación son los relacionados con la expresión de ER, PR, HER2 y el regulador de la proliferación celular Ki-67 (Tabla 1.1.1). Actualmente, el cáncer de mama se clasifica en los siguientes cinco subtipos

tumorales: luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo (10,19,21,22):

Subtipo luminal A: este subtipo representa el 50% de todos los cánceres de mama. Es el subtipo más común y menos agresivo, de muy buen pronóstico y con muy baja expresión de genes proliferativos. Se caracteriza por ser ER+, PR+, HER2- y tener una baja expresión de Ki-67. Al expresar ER, generalmente son susceptibles a tratamientos con hormonoterapia (tamoxifeno o inhibidores de la aromatasas), añadido al tratamiento quirúrgico o radio/quimioterápico que pueda requerir.

Subtipo luminal B HER2-: se caracteriza por ser ER+, PR+ o PR-, HER2- y tienen una expresión de Ki-67 intermedia. El aumento de expresión de Ki-67 hace que sean de crecimiento más rápido que los luminal A y que sean de peor pronóstico. Generalmente, son tratados con hormonoterapia junto con tratamiento quirúrgico y quimioterapia.

Subtipo luminal B HER2+: a diferencia del subtipo luminal B HER2-, este subtipo se caracteriza por ser ER+, PR+ o PR- y positivo para HER2. Tiene una expresión de Ki-67 intermedia-alta, lo cual hace que sea de crecimiento rápido y de peor pronóstico. Al ser HER2+ puede ser tratado con fármacos específicos de anticuerpos monoclonales dirigidos contra la proteína HER2/neu (Trastuzumab o Herceptin, Pertuzumab), además de con hormonoterapia, cirugía y tratamiento con quimioterapia.

Subtipo HER2+: este grupo representa el 15% de todos los cánceres de mama. Este subtipo se caracteriza por ser ER- y PR-, es positivo para HER2 y la expresión de Ki-67. Generalmente, suelen presentar una alta proporción de mutaciones en p53. Al no presentar receptores hormonales, no es susceptible al tratamiento hormonal, por lo que suele ser tratado con agentes específicos anti-HER2. La elevada expresión de HER2 y Ki-67, hace que sea un cáncer de crecimiento rápido. Es un subtipo de cáncer agresivo y de mal pronóstico.

Triple negativo: este subtipo comprende aproximadamente el 15% de todos los cánceres de mama. Se caracteriza por ser ER-, PR- y HER2-. Además, presentan una alta expresión Ki-67, mutaciones en p53 y son CK5/6 y EGFR positivos. Dado a que no expresa ER ni HER2, no responde al tratamiento hormonal ni a los agentes específicos anti-HER2, pero parece ser sensible a la quimioterapia a base de platino y a los inhibidores de la poli (ADP-ribosa) polimerasa (PARP). Es el subtipo más agresivo y, por lo general, tiene un pronóstico desfavorable.

Tabla 1.1.1. Características generales de los subtipos moleculares de cáncer de mama.

Subtipo tumoral	Expresión génica			Marcador de proliferación Ki-67
	Receptor de estrógeno (ER)	Receptor de progesterona (PR)	Receptor del factor de crecimiento epidérmico humano 2 (HER2)	
Luminal A	+	+	-	Baja
Luminal B HER-	+	+/-	-	Intermedia
Luminal B HER+	+	+/-	+	Intermedia-alta
HER2+	-	-	+	Alta
Triple negativo	-	-	-	Alta

1.2. MODIFICACIONES EN EL ARN

Durante décadas, el dogma central de la biología ha sido que la información fluía del ADN a las proteínas a través de la traducción. En este proceso, se consideró que el ARN era un participante intermedio que enviaba el código genético del cromosoma al ribosoma. Sin embargo, estudios emergentes han descubierto una gran cantidad de procesos biológicos en los que el ARN juega un papel regulador activo. En este sentido, se ha observado que el ARN puede sufrir modificaciones químicas. Este proceso, llamado epitranscriptómica, en analogía con la modificación postraduccional de proteínas y las modificaciones del ADN conocidas como epigenética, altera el comportamiento funcional de los ARN modificados.

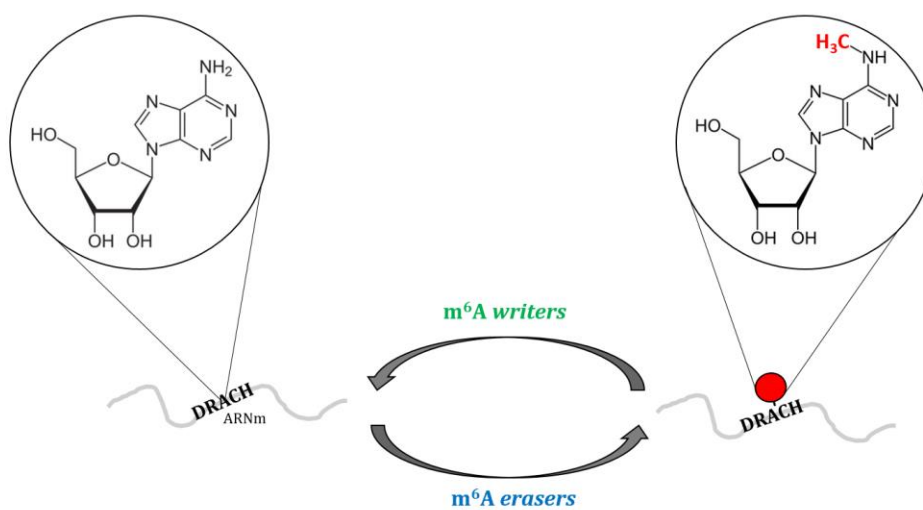
Hasta la fecha, se han identificado 163 modificaciones postranscripcionales en el ARN (23). Estas modificaciones pueden influir directamente en la estructura del ARN, promoviendo o interrumpiendo ciertas interacciones intramoleculares, haciendo que la molécula de ARN sea más rígida o más flexible, y también, pueden influir en las interacciones del ARN con otras moléculas, en particular con proteínas. En general, contribuyen fuertemente a la diversidad de funciones que cumplen las moléculas de ARN, donde pequeños cambios estructurales sutiles pueden provocar cambios significativos en el metabolismo celular (23,24).

1.2.1. Modificación m^6A

La N^6 -metiladenosina (m^6A) es una de las modificaciones más abundante en el ARN mensajero (ARNm) eucariota. Esta modificación fue descubierta en la década de 1.970 (25–30), sin embargo, su interés resurgió en el 2.011 con el descubrimiento de que la proteína asociada a la masa grasa y la obesidad (FTO) actúa como desmetilasa del ARN (31), lo que convierte a la modificación m^6A en un proceso dinámico y reversible (32).

El principal desafío para el análisis de las modificaciones m⁶A ha sido, durante mucho tiempo, la falta de una tecnología capaz de revelar las posiciones metiladas a lo largo del transcriptoma. Esto se debe a que el patrón de metilación m⁶A es eliminado durante la acción de la transcriptasa inversa en la reacción de PCR, y a que tampoco existe un tratamiento químico análogo a la conversión de bisulfito disponible para convertir el m⁶A en un nucleótido diferente y detectable. Sin embargo, en 2012, Meyer *et al.* (33) y Dominissini *et al.* (34) desarrollaron de forma independiente una tecnología que combinaba la inmunoprecipitación del ARN,

A



B

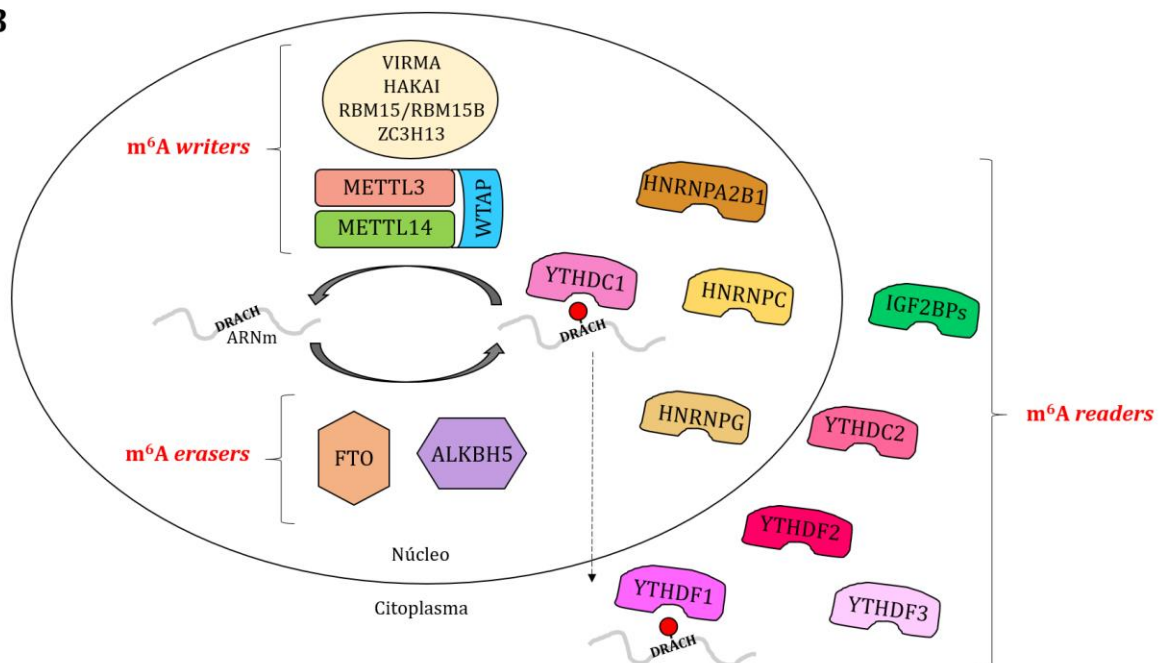


Figura 1.2.1. A- Proceso de metilación y desmetilación. **B-** Proteínas involucradas en la metilación m⁶A: *writers*, *erasers* y *readers*.

utilizando anticuerpos anti-m⁶A, seguida por la secuenciación masiva de alto rendimiento (MeRIP-seq), y que permitía mapear los sitios m⁶A en el transcriptoma de mamíferos. De estos estudios se concluyó que la modificación m⁶A ocurre con una frecuencia estimada de 3 a 5 sitios por transcripto en mamíferos, en regiones altamente conservadas con un motivo consenso identificado como RRA*CH (R = G o A; A*: adenina metilable; H = A, C o U), donde la adenina es un nucleótido que se metila en la posición N⁶ (Figura 1.2.1A). Posteriormente, el motivo consenso se redefinió como DRACH (D = G, A o U), basándose en el estudio de Linder *et al.* (2015) (35) donde informaron que el nucleótido en la posición -2 en relación con la posición A metilable también podría ser un uracilo. Además, se observó que las modificaciones m⁶A se encontraban enriquecidas en ciertas regiones del ARNm, como en la región 5' no traducida (UTR), en los exones largos, alrededor del codón de terminación y en la región 3'-UTR (33,34,36–38). También se ha observado que los motivos DRACH pueden estar presentes en el ARN no codificante (ARNnc) (37), microARNs (miARNs) (39), ARN ribosómico (ARNr) (34), ARN pequeño nuclear (ARNpn) (35) y ARN circular (40), aunque en menor proporción que en el ARNm.

1.2.2. Enzimas involucradas en el proceso de metilación/desmetilación

El proceso de metilación se encuentra regulado de forma dinámica y reversible por proteínas denominadas metiltransferasas o “*writers*”, desmetilasas o “*erasers*” y proteínas de unión a m⁶A o “*readers*” (Figura 1.2.1B). Las metiltransferasas son responsables de la metilación m⁶A en la posición N⁶ de la adenina, mientras que las desmetilasas eliminan la metilación m⁶A. Por otra parte, las proteínas de unión a m⁶A determinan el destino de los ARNm modificados al unirse al m⁶A. La expresión aberrante de estos componentes reguladores da como resultado la desregulación de m⁶A, lo cual se ha visto implicado en el inicio y la progresión de diversas enfermedades, incluido el cáncer (41).

1.2.2.1. Proteínas metiltransferasas o “*writers*”

Este grupo de proteínas metiltransferasas o “*escritoras*” está formado por el complejo de metiltransferasas (METTL3, METTL14 y WTAP) y varios cofactores (VIRMA, HAKAI, RBM15/RBM15B y ZC3H13) (42–44).

La proteína metiltransferasa 3 (METTL3) fue la primera subunidad del complejo de metiltransferasas identificada, en 1.977, y es la única que presenta actividad catalítica (45). METTL3 contiene un dominio de unión a S-adenosilmetionina (SAM) y un motivo DPPW (Asn-Pro-Pro-Trp) que transfiere el grupo metilo desde SAM a la posición N⁶ de la adenina. Aunque METTL3 es la única subunidad catalítica del complejo de metiltransferasas, se encuentra inactiva sin la subunidad formada por la proteína metiltransferasa 14 (METTL14). Los estudios bioquímicos mostraron que ambas subunidades forman un heterodímero, en donde METTL14 reconoce los sustratos de ARN, sirve como plataforma de unión al ARN, facilita la activación alostérica y promueve la actividad catalítica de METTL3, quien es la responsable de generar la metilación m⁶A como tal (46–48). La tercera subunidad del complejo de metiltransferasas es la proteína asociada al tumor 1 de Wilms (WTAP). Esta proteína interactúa con METTL3-METTL14 promoviendo la localización de este heterodímero en las motas nucleares, es decir, sitios de empalme y almacenamiento de ARNm y facilitando la metilación de los ARNm (49).

De igual modo, diversos cofactores se unen al complejo de metiltransferasas, guiándolos a sitios específicos en el ARN y regulando el proceso de metilación. En este sentido, el homólogo del virilizador de proteínas (VIRMA, también conocido como KIAA1429) media la metilación preferencial del ARNm en la región 3'-UTR y cerca del codón de parada (50). La proteína de unión al ARN 15 (RBM15) y su paróloga (RBM15B) reconocen secuencias ricas de uracilos en el ARNm y guían al complejo de metiltransferasas a zonas adyacentes a estos sitios (51). La proteína E3 ubiquitina ligasa Hakai (HAKAI) estabiliza a los componentes centrales de la maquinaria de metilación m⁶A (52), mientras que la proteína de dedo de zinc ZC3H13 regula la metilación m⁶A.

1.2.2.2. Proteínas desmetilasas o “erasers”

Aunque METTL3 fue descubierta como una proteína “*writer*” varias décadas atrás, la identidad de las proteínas desmetilasas o “*erasers*” siguió siendo un misterio hasta 2.011, cuando Jia *et al.* (31) revelaron que la proteína FTO exhibía una actividad desmetilasa de m⁶A. La demostración de su actividad desmetilasa proporcionó la primera evidencia de que la metilación m⁶A era una modificación reversible. FTO también muestra actividad desmetilasa hacia N⁶-dimetiladenosina (m⁶Am) que se encuentra exclusivamente en el primer nucleótido codificado después de la estructura de cap-7-metilguanosa de un gran número de ARNm, por lo que su actividad desmetilasa no es específica para m⁶A (53). En este sentido, se ha observado que FTO

presenta una baja tasa de desmetilación para el m⁶A en comparación con el m⁶Am (54) y que en el núcleo celular FTO media principalmente la desmetilación de m⁶A, mientras que en el citoplasma puede mediar la desmetilación de m⁶A así como la de m⁶Am (55).

En el 2013, se descubrió una segunda proteína desmetilasa de m⁶A, el homólogo 5 de AlkB (ALKBH5) (56) que pertenece a la familia AlkB de dioxigenasas dependientes de Fe (II)/ α -cetoglutarato, al igual que FTO. Sin embargo, aunque ambas desmetilasas pertenecen a la misma familia, presentan diferencias en cuanto a la desmetilación: FTO desmetila m⁶A a través de tres rondas de oxidación para producir dos productos intermedios (N⁶-hidroximetiladenosina y N⁶-formiladenosina), mientras que ALKBH5 convierte directamente el m⁶A en adenosina sin productos intermedios observados (57). Además, ALKBH5 no tiene actividad hacia m⁶Am, por lo que es específica para m⁶A (53).

1.2.2.3. *Proteínas de unión o “readers”*

Las proteínas de unión también llamadas “readers”, reconocen y se unen preferentemente al ARN que contiene la modificación m⁶A. Este grupo se compone principalmente de proteínas que contienen el dominio YTH (YTHDF1-3 y YTHDC1-2), ribonucleoproteínas nucleares heterogéneas (hnRNPC, hnRNPG y hnRNPA2B1) y proteínas de unión al ARNm del factor de crecimiento similar a la insulina tipo 2 (IGF2BPs) (58).

De acuerdo a su distribución subcelular, la familia de dominios YTH se clasifica en proteínas de la familia de dominios YTH (YTHDF) compuesto por YTHDF1, YTHDF2 y YTHDF3, y proteínas que contienen el dominio YTH (YTHDC), siendo, en este caso YTHDC1 y YTHDC2 (59,60). En el núcleo, la proteína YTHDC1 interactúa con el m⁶A en el ARNm para regular el empalme alternativo, la poliadenilación del pre-ARNm y la exportación nuclear (59–61). Por otra parte, en el citoplasma, YTHDF1 interactúa con factores de iniciación, como por ejemplo eIF3, para facilitar la iniciación de la traducción del ARNm. YTHDF2 se une selectivamente al ARNm m⁶A metilado, reduciendo su estabilidad y regulando su degradación. YTHDF3 facilita la traducción al promover la síntesis de proteínas en sinergia con YTHDF1 y afecta la desintegración del ARNm mediada por YTHDF2. Por lo que, las tres proteínas YTHDF funcionan de manera cooperativa (59,62). Por otra parte, se ha visto que la proteína YTHDC2 aumenta la eficiencia de la traducción de los ARNm metilados (63).

Miembros de la familia HNRNP también tienen el potencial de identificar la modificación m⁶A de los ARNm. En este sentido, se ha visto que las proteínas HNRNPC y HNRNPG regulan el procesamiento de los pre-ARNm que contienen m⁶A (42), mientras que la proteína HNRNPA2B1 acelera el procesamiento de miARN primarios (pri-miARN) y modula el empalme alternativo de las transcripciones (64,65). Otro grupo de *readers* son las proteínas de unión al ARNm del factor de crecimiento similar a la insulina tipo 2 (IGF2BPs). Estas proteínas estabilizan las transcripciones dianas que contienen la metilación m⁶A y promueven la traducción (66).

1.2.3. Funciones biológicas de la modificación m⁶A

La modificación m⁶A está involucrada en la regulación de la mayoría de los procesos del metabolismo del ARN (67). En el núcleo, la modificación m⁶A está involucrada en la regulación del procesamiento del pre-ARNm (64), empalme alternativo del ARNm (61,68–70), la exportación del ARNm del núcleo al citoplasma (71) y la poliadenilación alternativa (38). En el citoplasma, la metilación m⁶A regula la traducción del ARNm (36,72,73) y afecta la estabilidad y desintegración del mismo (66,74). Además, la modificación m⁶A cambia la estructura del ARN alterando también las interacciones ARN-proteínas (68,75,76). Por otra parte, en los miARNs, la modificación m⁶A puede alterar el procesamiento de los pri-miARNs (77) y regular la maduración de los miARNs (65).

Las funciones de las modificaciones del ARN en los procesos biológicos acaban de empezar a apreciarse. Se ha visto que la modificación m⁶A se encontraba implicada, por ejemplo, en el mantenimiento y la diferenciación de las células madre embrionarias (78–80), la modificación del ritmo circadiano (81), la respuesta al choque térmico (82) y progresión meiótica (83). Sin embargo, la función que poseen en la mayoría de los procesos biológicos y enfermedades sigue siendo desconocida o incompleta. Específicamente, aún queda por determinar todas las funciones que la modificación m⁶A presenta en el cáncer de mama.

1.2.4. Métodos de análisis del m⁶A

Son diversos los métodos que se han desarrollado, y continúan desarrollándose, con el fin de analizar las posiciones m⁶A. La Tabla 1.2.1 describe los principales métodos utilizados para determinar las posiciones m⁶A metiladas en todo el transcriptoma y aquellos que permiten su

cuantificación. La gran mayoría de las técnicas se diferencian en si utilizan anticuerpos u otros compuestos para identificar las posiciones m⁶A y si pueden ser empleadas en ARN procedentes de tejidos y cultivos celulares o únicamente de cultivos celulares. Hasta la fecha, los métodos más utilizados para la identificación de las posiciones m⁶A han sido MeRIP-seq, MiCLIP y PA-m⁶A-seq. Por lo que, a continuación, se describen brevemente estas técnicas.

- MeRIP-seq

La inmunoprecipitación del m⁶A-ARN, seguida de la secuenciación de alto rendimiento de próxima generación (NGS, *Next Generation Sequencing*) más comúnmente conocida como MeRIP-seq, fue el primer método desarrollado para identificar las modificaciones m⁶A en todo el transcriptoma (33,34) (Figura 1.2.2). Hasta la fecha, continúa siendo el método más empleado para dicho fin (84). En resumen, esta técnica consiste en inmunoprecipitar con un anticuerpo anti-m⁶A el ARN previamente fragmentado (85). Posteriormente, se genera la librería de los fragmentos inmunoprecipitados (muestra *IP*), así como de la muestra de ARN fragmentada sin inmunoprecipitar (muestra *INPUT*) y se secuencian mediante NGS. Finalmente, las secuencias se alinean y se realiza la determinación de las posiciones m⁶A metiladas mediante diversos programas informáticos, como por ejemplo MACS (86), exomePeak (87) y MeTPeak (88). Esta técnica no permite identificar la ubicación precisa de los sitios m⁶A dentro de un fragmento de ARN inmunoprecipitado, sino que presenta una resolución de unos 100-200 nucleótidos (nts). Esto puede dificultar la identificación de las posiciones m⁶A agrupadas, es decir, posiciones m⁶A que se encuentran próximas en el mismo fragmento inmunoprecipitado. Otra limitación de la técnica es la cantidad inicial de ARN de partida. En relación a esto, Zeng *et al.* (2018) (89) han optimizado el protocolo logrando obtener perfiles de metilación m⁶A utilizando 2 µg de ARN total (89). Sin embargo, a pesar de estas limitaciones, MeRIP-seq cuenta con las ventajas de que es un protocolo simple y corto, y que utiliza materiales, equipos y paquetes de software accesibles (85).

Tabla 1.2.1. Métodos utilizados para la detección y la cuantificación de las posiciones m⁶A metiladas.					
Métodos	Detección de m⁶A	Ventajas	Desventajas	Ref.	
Detección de las posiciones m ⁶ A metiladas	MeRIP-seq	Anticuerpo	-Protocolo simple y corto -Utiliza materiales, equipos y paquetes de software accesibles	-Resolución de unos 100-200 nts -Dificultad al identificar posiciones m ⁶ A agrupadas	(85,89)
	MiCLIP	Anticuerpo	-Resolución de un nucleótido	-Técnica larga, laboriosa y cara -Las firmas mutacionales dependen de los anticuerpos anti-m ⁶ A utilizados -Utilización de ³² P	(90-92)
	PA-m ⁶ A-seq	Anticuerpo	-Resolución de 23 nts	-Uso en muestras procedentes de cultivos celulares, ya que previamente las células deben incorporar 4-tiouridina (s ⁴ U) -Detecta las posiciones que hayan incorporado s ⁴ U	(93)
	DART-seq	APOBEC1-YHD	-Utiliza 10 nanogramos de ARN celular total -Compatible con cualquier método de preparación de bibliotecas de ARN-seq estándar -Reduce el tiempo y el coste en comparación con las técnicas basadas en anticuerpos	-Uso en muestras procedentes de cultivos celulares, ya que previamente las células deben incorporar el plásmido APOBEC1-YHD -No hay kits comerciales que proporcionen el plásmido APOBEC1-YHD	(94)
	Secuenciación por nanoporo		-Permite la secuenciación directa de moléculas de ARN de longitud completa y, al mismo tiempo, obtener información sobre las modificaciones presentes en el ARN	-Permite la secuenciación directa de moléculas de ARN de longitud completa y, al mismo tiempo, obtener información sobre las modificaciones presentes en el ARN	(95-97)
	m ⁶ A-REF-seq	Toxina MazF	-Rápido, simple, no necesita un paso de enriquecimiento de m ⁶ A, lo cual reduce la cantidad de ARN de partida y el tiempo de preparación de las muestras -Resolución de un nucleótido	-Únicamente identifica las posiciones m ⁶ A presentes en el motivo DRACA -MazF no puede escindir en el ARN bicatenario, por lo que, no identificará la presencia de m ⁶ A en el ARN estructurado	(98)
	m ⁶ ACE-seq	Exorribonucleasa XRN1	-Resolución de un nucleótido		(99)

Métodos	Detección de m ⁶ A	Ventajas	Desventajas	Ref.
Detección de las posiciones m ⁶ A metiladas	m ⁶ A-SEAL Enzima FTO	<ul style="list-style-type: none"> -Utiliza reactivos disponibles comercialmente, -Buena sensibilidad, especificidad y fiabilidad -Puede utilizarse con baja cantidad de 	-Resolución de 200 nts.	(100)
	m ⁶ A-label-seq Anticuerpo	<ul style="list-style-type: none"> -Permite identificar las posiciones m⁶A agrupadas 	-Uso en muestras procedentes de cultivos celulares, ya que previamente las células deben incorporar	(101)
Semicuantitativos y Cuantitativos	Dot blot Anticuerpo	-Reduce el tiempo de análisis ya que no requiere cromatografía o electroforesis en gel	<ul style="list-style-type: none"> -Solo puede verificar la presencia o ausencia del m⁶A o comparar las cantidades de m⁶A entre diferentes grupos -No ofrece información sobre el tamaño del ARN metilado, y, por ende, de qué tipo de ARN se trata. 	(102,103)
	Immuno-northern blot Anticuerpo	<ul style="list-style-type: none"> -Técnica simple -Separa los ARN según sus pesos moleculares, permitiendo detectar la modificación m⁶A en diferentes tipos de ARN 	<ul style="list-style-type: none"> -Técnica laboriosa -Baja sensibilidad -Requiere ARN de muy buena calidad 	(104,105)
Semicuantitativos y Cuantitativos	LAIC-seq Anticuerpo	<ul style="list-style-type: none"> - Es posible comparar isoformas empalmadas alternativamente 	<ul style="list-style-type: none"> -Los resultados del análisis dependen en gran medida del algoritmo bioinformático utilizado para llamar a los picos 	(106)
	MAZTER-seq Toxina MazF	-Genera datos estequiométricos	<ul style="list-style-type: none"> -Permite la cuantificación de sitios m⁶A que ocurren en el motivo DRACA -MazF también podría escindir aquellos motivos que se asemejan a ACA, como ACG o AAA, lo cual puede sesgar las cuantificaciones -Dificultades para comparar los resultados obtenidos entre muestras 	(107,108)

Continuación de la Tabla 1.2.1.

- MiCLIP

Este método permite la identificación de posiciones m⁶A a resolución de un nucleótido, mediante la unión covalente del anticuerpo con la posición m⁶A metilada presente en el ARN (*crosslinking*) y la inmunoprecipitación (90) (Figura 1.2.2). En este sentido, los anticuerpos anti-m⁶A se enlazan covalentemente con el ARN en la posición metilada usando luz UV para crear firmas mutacionales específicas que permiten la identificación precisa de residuos de m⁶A. Estas firmas mutacionales pueden dar lugar a una transición de base o un truncamiento de la secuencia, que podrán ser detectadas en el ADN complementario (ADNc) tras realizar la transcripción inversa. Finalmente, las librerías son secuenciadas y las firmas mutacionales se identifican computacionalmente, revelando las posiciones precisas de los residuos de m⁶A (91,92). A pesar de la capacidad que posee esta técnica de identificar la posición exacta de la metilación m⁶A, cuenta con algunas desventajas. Por un lado, es una técnica larga, laboriosa y cara. Por otro lado, la capacidad de generar las firmas mutacionales depende de los anticuerpos utilizados. Generalmente, los anticuerpos más utilizados son los anticuerpos anti-m⁶A de la casa

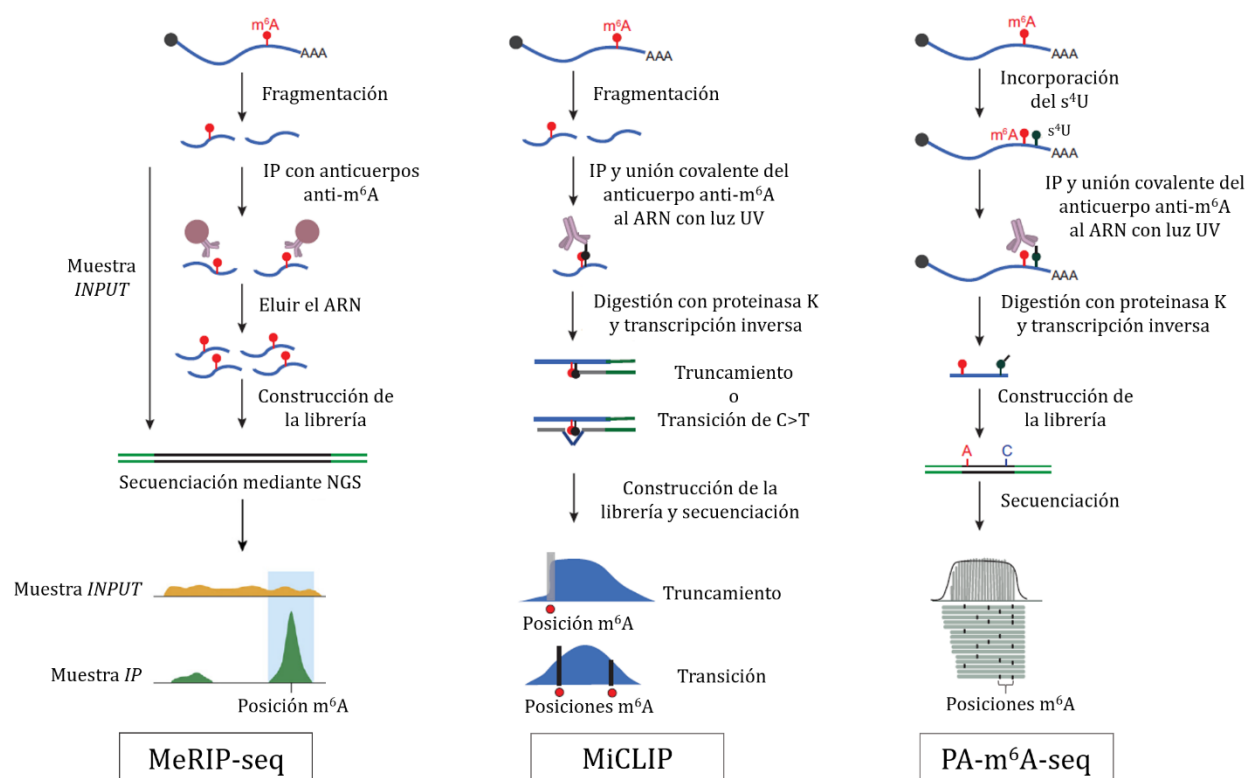


Figura 1.2.2. Métodos utilizados para la detección de la modificación m⁶A: MeRIP, MiCLIP y PA-m⁶A-seq. Figura modificada de Li *et al.* (2.016) (109). IP: inmunoprecipitación.

comercial *ABCAM* (ab151230) que producen una transición de la citosina del motivo DRACH por una timina, y los anticuerpos anti-m⁶A de la casa comercial *Synaptic System* (202003), que producen un truncamiento de la secuencia en la citosina del motivo DRACH (90). Otra desventaja es que uno de los pasos consiste en un Western blot y transferencia a membrana utilizando fósforo radiactivo (³²P). Sin embargo, esta desventaja se podría eliminar al sustituir el ³²P por un adaptador conjugado con un colorante infrarrojo, tal y como describen Zarnegar *et al.* (2016) (110).

- PA-m⁶A-seq

La técnica PA-m⁶A-seq permite identificar las posiciones m⁶A con una resolución de 23 nts (93) (Figura 1.2.2). En esta técnica al añadir 4-tiouridina (s⁴U) al medio de cultivo, las células en cultivo captan e incorporan el s⁴U al ARN cerca de las posiciones m⁶A. Posteriormente, se realiza la inmunoprecipitación del ARN sin fragmentar y se irradia con luz UV de 365 nm para producir la unión covalente entre el anticuerpo y el ARN. El ARN unido al anticuerpo se fragmenta hasta 25-30 nts y se trata con proteinasa K para liberar el ARN del anticuerpo. Los fragmentos de péptidos que permanecen después de la eliminación de anticuerpos del ARN hacen que la citosina sea sustituida por una timina durante la transcripción inversa, lo que induce a una mutación puntual en el sitio s⁴U. Finalmente, se preparan las librerías y se secuencian (93). Esta técnica cuenta con las desventajas de que únicamente se puede realizar en muestras procedentes de cultivos celulares ya que previamente tienen que ser tratadas con s⁴U. Además, aquellos sitios que no hayan incorporado s⁴U no podrían ser detectados.

1.3. LA METILACIÓN m⁶A Y EL CÁNCER DE MAMA

En los últimos años, diversas investigaciones han concluido que la aparición, desarrollo y progresión del cáncer son consecuencia no sólo de alteraciones genéticas, sino también de modificaciones epigenéticas. En este sentido, se ha observado que las modificaciones epigenéticas en el ADN y las modificaciones en las histonas estaban relacionadas con el cáncer de mama (111–117). Recientemente se ha descubierto que las modificaciones químicas en el ARN también podían estar involucrados en el cáncer (118–120). Por lo que, la modificación m⁶A en el ARNm abre una nueva era para el estudio del desarrollo y la progresión, como también, la búsqueda de nuevas dianas terapéuticas en el cáncer de mama (121–123).

1.3.1. Metiltransferasas, desmetilasas y proteínas de unión al m⁶A en el cáncer de mama

Dada la importancia que tiene la modificación m⁶A en los aspectos metabólicos del ARN, como también en distintos procesos biológicos, la expresión aberrante se encuentra estrechamente relacionada con el desarrollo y mantenimiento de diversas enfermedades, incluido el cáncer (41). En este sentido, las metiltransferasas y desmetilasas cambian los niveles de m⁶A influyendo en el empalme, desintegración, exportación nuclear del ARNm y la traducción para alterar la expresión del gen diana, lo que impacta en los fenotipos del cáncer. Además, las proteínas de unión al m⁶A detectan sitios m⁶A y ejercen funciones relativas (124).

Diversos estudios se han llevado a cabo en el cáncer de mama con el fin de determinar la expresión de los genes que están involucrados en la modificación m⁶A como se muestra en la Tabla 1.3.1. Se puede observar la expresión alterada de los modificadores m⁶A en comparación con la expresión en el tejido sano. Hay que señalar que la gran mayoría de estos estudios se han realizado *in silico*, utilizando los datos presentes en distintas bases de datos como ONCOMINE, TCGA y GEO. Únicamente en los trabajos realizados por Wu *et al.* (2.019) (125) y Yi *et al.* (2.020) (126) se ha llevado a cabo el análisis de expresión *in vitro* en muestras de tejido tumoral y sano de pacientes de China.

Tabla 1.3.1. Resultados de los estudios llevados a cabo en el cáncer de mama sobre la expresión de los genes involucrados en la modificación m⁶A. Se detallan las muestras analizadas en cada estudio, como también el estado de la expresión, es decir, si la expresión es alta o baja en comparación con el tejido normal.

Muestras analizadas	Expresión génica alta	Expresión génica baja	Ref.
Muestras de la base de datos ONCOMINE	<i>ALKBH5</i>	<i>METTL3, METTL14, WTAP y FTO</i>	(125)
Muestras de la base de dato TCGA	-	<i>METTL3, METTL14, WTAP, FTO y ALKBH5</i>	
Muestras tumorales y sanas de 36 pacientes procedentes de China	<i>ALKBH5</i>	<i>METTL3, METTL14, WTAP y FTO</i>	
111 muestras tumorales y 12 muestras sanas procedentes de la base de datos GEO	<i>FTO</i>	<i>METTL3, METTL14, YTHDC2, YTHDF1 y ALKBH5</i>	(127)
Seis datasets obtenidos de la base de datos TCGA	<i>VIRMA, HNRNPA2B, HNRNPC, IGF2BP1, YTHDF1 y ZNF217</i>	<i>METTL14, WTAP, IGF2BP2, YTHDC1 y FTO</i>	(128)
1.066 muestras tumorales y 112 muestras sanas procedentes de la base de datos TCGA	<i>VIRMA, HNRNPC, YTHDF1, RBM15, YTHDF2 y HNRNPA2B1</i>	<i>METTL14, WTAP, YTHDF3, YTHDC1, FTO y ALKBH5</i>	(129)
Muestras procedentes de la base de datos GEO	<i>VIRMA, RBM15, YTHDF1, YTHDF2, HNRNPC, IGFBP2 y HNRNPA2B1</i>	<i>METTL14, RBM15B, YTHDC1, IGFBP1, IGFBP3 y FTO</i>	(130)
Muestras tumorales y sanas de 20 pacientes procedentes de China	<i>METTL14</i>	-	(126)

Muestras analizadas	Expresión génica alta	Expresión génica baja	Ref.
Muestras de la base de datos ONCOMINE	-	<i>METTL3, METTL14 y RMB15B</i>	(131)
1.091 muestras tumorales y 113 muestras sanas procedentes de la base de datos TCGA	<i>IGF2BP1, IGF2BP3, RBM15, KIAA1429, YTHDF2, YTHDF1 y HNRNPA2B1</i>	<i>IGF2BP2, METTL14, YTHDC1 y WTAP</i>	(132)

Además del análisis de expresión génica de los modificadores de m⁶A, diferentes estudios han demostrado que las proteínas METTL3, METTL14, FTO, ALKBH5, YTHDF3 e IGF2BP1 ejercen funciones dependientes de la modificación m⁶A en el cáncer de mama (133). La Tabla 1.3.2 recoge los estudios llevados a cabo sobre estos modificadores de m⁶A, así como, la/s molécula/s diana/s y eje de señalización donde actúan y el efecto que producen en el cáncer de mama.

Tabla 1.3.2. Funciones de las proteínas modificadoras de m⁶A en el cáncer de mama. Modificado de Wang *et al.* (2020) (120).

Rol en la modificación m ⁶ A	Proteína	Diana/Eje de señalización	Función biológica	Ref.
Metiltransferasas	METTL3	Bcl-2	Promover la proliferación celular, el crecimiento tumoral e inhibir la apoptosis celular	(134)
		HBXIP/let7-g /METTL3/HBXIP	Promover la proliferación celular	(135)
		LINC00958	Promover la progresión tumoral de las células del cáncer de mama	(136)
		COL3A1	Promover la metástasis	(137)
	METTL14	Has-miR-146a-5p	Promover la migración e invasión de las células de cáncer de mama	(138)
		LNC942-METTL14-CXCR4/CYP1B1	Promover la proliferación celular, la formación de colonias e inhibir la apoptosis celular	(139)
Desmetilasas	FTO	BNIP3	Promover la proliferación celular, la formación de colonias y la metástasis.	(127)
	ALKBH5	NANOG	Aumentar el porcentaje de células madre del cáncer de mama	(140)
		NANOG y KLF4	Aumentar el porcentaje de células madre del cáncer de mama	(141)
Proteínas de unión al m ⁶ A	YTHDF3	ST6GALNAC5, GJA1 y EGFR	Promover la metástasis cerebral del cáncer de mama	(142)
	IGF2BP1	lncRNA KB-1980E6.3/ IGF2BP1/c-Myc	Promover la autorrenovación y tumorigénesis de las células madre del cáncer de mama	(143)

1.3.2. *m*⁶A asociados a SNP (*m*⁶A-SNP)

Estudios recientes han indicado que los polimorfismos de un solo nucleótido (*Single Nucleotide Polymorphisms*, SNP por sus siglas en inglés) ubicados dentro o cerca de los motivos de DRACH pueden afectar la metilación *m*⁶A y los procesos biológicos relacionados al alterar la secuencia del ARN o los nucleótidos flanqueantes (144,145). Estas variaciones genéticas se conocen como *m*⁶A asociados a SNP (*m*⁶A-SNP).

La importancia de estas variantes se debe a que los *m*⁶A-SNPs podrían tener un efecto de *loci* de rasgos cuantitativos de expresión (*Expresión Quantitative Trait Loci*, eQTL por sus siglas en inglés), es decir, podrían modificar los niveles de expresión génica y, en consecuencia, influir en el desarrollo y progresión enfermedades (146). Específicamente, los *m*⁶A-SNPs podrían tener un efecto *cis*-eQTL afectando la expresión de genes cercanos a la posición en que se encuentran. Este efecto puede deberse a múltiples causas como la alteración del nivel de metilación *m*⁶A y de la unión de factores de transcripción y proteínas, la estabilidad del ARNm, la exportación nuclear, el alargamiento de la transcripción, el empalme o la modificación de la estructura secundaria del ARNm (72,144,147,148).

Generalmente, todos los estudios que identifican *m*⁶A-SNPs relacionados a distintas enfermedades han sido realizados *in silico*, utilizando los datos disponibles en distintas bases de datos. Los estudios llevados a cabo por Chen *et al.* (2021) en diabetes tipo 2 (149) y Ruan *et al.* (2021) en el cáncer de tiroides (150), fueron los únicos que validaron los *m*⁶A-SNPs identificados mediante el análisis de la expresión génica.

Hasta la fecha, se ha identificado el papel de los *m*⁶A-SNPs en varias enfermedades, como la artritis reumatoide (151), osteoporosis (152,153), la enfermedad de las arterias coronarias (154), la periodontitis (147), el desarrollo de sepsis (148), la adiposidad (155), niveles de lípidos (colesterol y triglicéridos) (156), enfermedad de Parkinson (157), la enfermedad inflamatoria intestinal (158), el cáncer colorrectal (159), diabetes tipo 2 (149) y el cáncer de tiroides (150). Recientemente, Xuan *et al.* (2021) (160) han encontrado tras el análisis *in silico*, ocho *m*⁶A-SNPs que podrían tener un efecto sobre el cáncer de mama. Estos *m*⁶A-SNPs se encontraban en los genes *TOM1L1* (rs1802212 y rs4829), *ESR1* (rs2747655), *MAFF* (rs2267372 y rs9610915), *TNS1* (rs1424916), *EFEMP2* (rs2234458) y *TRIM31* (rs2023472). Estos resultados muestran únicamente el inicio del potencial que podrían tener los *m*⁶A-SNPs en el cáncer de mama. Por lo que, existe la necesidad de realizar más estudios utilizando los datos disponibles en otras bases de datos con el fin de tener una comprensión más completa del papel

que podrían desempeñar los m⁶A-SNPs en el cáncer de mama, además de realizar estudios *in vitro* para evaluar sus posibles efectos sobre la expresión génica.

1.3.3. Perfil de metilación m⁶A en todo el transcriptoma

Las últimas investigaciones llevadas a cabo en la modificación m⁶A se han centrado en identificar el perfil de metilación m⁶A presente en todo el transcriptoma, es decir, el conjunto de moléculas de ARNm y de ARNnc presentes en un determinado tejido. Este tipo de análisis, además de identificar las posiciones m⁶A metiladas, permite visualizar qué regiones y transcritos se encuentran enriquecidos con m⁶A revelando su potencial efecto, así como posibles dianas terapéuticas.

Hasta la fecha, se ha identificado el perfil de metilación m⁶A presente en muestras relacionadas con el cáncer de vejiga (161), carcinoma de células renales de células claras (162), adenoma pleomórfico invasivo (163), ameloblastoma, (164) trastornos del espectro de la neuromielitis óptica (165), alta miopía (166), insuficiencia cardíaca (167), osteosarcoma (168) y preeclampsia (169). Todas estas investigaciones se han realizado utilizando la técnica MeRIP-seq y analizando el ARN procedente de tejidos asociados a dichas enfermedades. Esto supone un gran avance, ya que, los datos obtenidos en estudios previos se habían realizado con ARN procedentes de cultivos celulares. En cuanto al cáncer de mama, sin embargo, el perfil de metilación m⁶A aún no ha sido identificado. Por ello, determinar el perfil de metilación m⁶A podría complementar la caracterización de los distintos subtipos moleculares, revelar una mejor predicción del comportamiento del tumor e identificar posibles estrategias terapéuticas específicas para cada subtipo.

2. HIPÓTESIS Y OBJETIVOS

2.1. HIPÓTESIS

La N⁶-metiladenosina (m⁶A) es una de las modificaciones más abundante del ARNm presente en los organismos eucariotas. Aunque esta modificación química se conoce desde 1.970, no ha suscitado gran interés hasta la última década. Esto puede deberse en gran medida al descubrimiento de enzimas capaces de revertir la metilación m⁶A, como también, al desarrollo de metodologías que permitieron identificar las posiciones m⁶A metiladas presentes en todo el transcriptoma. Varios estudios han revelado que la modificación m⁶A está involucrada en la regulación de la mayoría de los procesos del metabolismo del ARN. Sin embargo, aún no existe mucho conocimiento acerca del impacto que esta modificación tiene sobre la regulación de la expresión génica, su función en los procesos metabólicos y el efecto que puede ejercer en el desarrollo y progresión de ciertas enfermedades.

Hoy en día, el cáncer de mama constituye la primera causa de muerte por cáncer en mujeres. El cáncer de mama presenta una elevada heterogeneidad biológica, histológica y clínica que dificulta la selección de un tratamiento adecuado. En este sentido, aunque la aplicación de nuevos tratamientos está consiguiendo aumentar la supervivencia de las pacientes, la búsqueda de nuevos fármacos todavía representa un objetivo principal de la terapia del cáncer. Por tal motivo, la hipótesis de partida del presente estudio es que la modificación m⁶A podría estar implicada en el cáncer de mama. La metilación m⁶A puede resultar susceptible a cambios en su estado de metilación a lo largo de la vida de un individuo y, en consecuencia, es muy probable que haya puntos en el transcriptoma cuyas variaciones en su estado de metilación m⁶A se encuentren relacionadas con el desarrollo y progresión del cáncer. En este sentido, la identificación de estas posiciones que varían en su estado de metilación m⁶A permitiría la búsqueda de posiciones diana para fármacos ya existentes que pudieran anular dichos puntos inestables de metilación o actuar como dianas para el diseño de nuevas moléculas terapéuticas que sirvan para revertir y/o frenar la progresión de la enfermedad. Además, dado que el cáncer de mama se clasifica en los subtipos moleculares luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo, la determinación de perfiles de metilación m⁶A permitiría aumentar su caracterización epigenética y la identificación de dianas terapéuticas propias de cada subtipo.

2.2. OBJETIVOS

El objetivo general de este estudio es explorar la presencia de la modificación m⁶A en el cáncer de mama, específicamente en los subtipos moleculares luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo, con el fin de determinar el estado de la metilación y de su posible relación con futuras dianas terapéuticas.

Con este fin, los objetivos específicos del trabajo fueron los siguientes:

1. Revelar el estado de la metilación m⁶A en muestras de pacientes con cáncer de mama, de manera general y teniendo en cuenta los distintos subtipos moleculares, mediante la determinación del porcentaje de metilación m⁶A global.
2. Evaluar el estado de expresión génica de las metiltransferasas *METTL3*, *METTL14* y *WTAP* y de las desmetilasas *FTO* y *ALKBH5*.
3. Identificar los m⁶A-SNPs asociados al cáncer de mama y evaluar sus posibles efectos sobre la expresión génica.
4. Determinar el perfil de metilación m⁶A presente en el transcriptoma de los distintos subtipos moleculares mediante el uso de la técnica MeRIP-seq.

3. MATERIALES Y MÉTODOS

3.1. MUESTRAS BIOLÓGICAS

En el presente estudio se han analizado muestras de tejido mamario congelado de 35 mujeres diagnosticadas con IBC-NST. Por cada paciente se obtuvo una muestra del tejido tumoral y otra muestra correspondiente al tejido mamario sano. Las muestras tumorales se clasificaron según su subtipo molecular: siete pertenecían al subtipo luminal A, siete al subtipo luminal B HER2-, ocho al subtipo luminal B HER2+, seis al subtipo HER2+ y siete al subtipo triple negativo. Además, se obtuvieron los siguientes datos asociados: edad, estado menopáusico al diagnóstico, recidivas, TNM, grado histológico, vascular linfática, multicentricidad del tumor, necrosis tumoral, receptores estrógeno y progesterona, HER2 (IHQ y FISH), Ki-67, CK19 y CK 5/6.

Tanto las muestras de tejido mamario congelado como los datos asociados a las mismas se obtuvieron del Biobanco Vasco tras conseguir la aprobación del Comité de Ética en Investigación de Medicamentos del País Vasco (PI2018044).

3.2. EXTRACCIÓN Y CUANTIFICACIÓN DE ARN TOTAL

3.2.1. Extracción de ARN total

La extracción del ARN total de las muestras tumorales y del tejido mamario sano se realizó a partir de aproximadamente 100 mg de tejido congelado. Tras obtener la cantidad necesaria, se procedió a la disgregación mecánica del tejido en nitrógeno líquido. La extracción del ARN total se llevó a cabo utilizando el kit PureLink® RNA Mini Kit (cat. N°: 12183025, Invitrogen, Carlsbad, USA) (en adelante PureLink®) en conjunto con el Reactivo TRIzol® (cat. N°: 15596018, Invitrogen, Carlsbad, USA) e incluyendo el tratamiento por DNasas PureLink® DNase Set (cat. N°: 12185010, Invitrogen, Carlsbad, USA), con algunas modificaciones. El protocolo utilizado para la extracción del ARN total fue el siguiente:

1. Primeramente, se descongeló el tejido durante el tiempo necesario para poder realizar el corte y se acondicionó la superficie para el corte del fragmento mediante una pieza de papel de aluminio doblada sobre sí misma al menos cuatro veces. Posteriormente, con la ayuda de una aguja hipodérmica estéril, se extrajo la pieza de tejido a cortar del tubo donde se encontraba almacenada, se depositó en el papel plegado de aluminio y se cortaron aproximadamente 100 mg de tejido con una hoja de bisturí estéril.

2. El fragmento de tejido cortado se envolvió con el papel de aluminio y se introdujo en un recipiente que contenía nitrógeno líquido durante 1 minuto. Pasado este tiempo, se sacó del nitrógeno líquido y con la ayuda de una maza, se dieron unos golpes al fragmento envuelto en papel de aluminio para disgregarlo hasta reducirlo a polvo. Este proceso (sumergir en nitrógeno líquido y golpear el fragmento de tejido con una maza) se repitió una vez más y, finalmente, se traspasó la muestra completamente pulverizada a un tubo de 1,50 mL libre de ARNasas (cat. N°: AM12400, Ambion, Carlsbad, USA).
3. Se incorporó 1 mL del reactivo TRIzol® y se homogenizó con el motor inalámbrico Kontes Pellet Pestle (Cat N°: 749541-0000, DWK Life Sciences, Mainz, Germany).
4. Se incubó 5 minutos a temperatura ambiente para permitir la completa disociación de los complejos nucleoproteicos.
5. Posteriormente, se centrifugó durante 5 minutos a $12.000 \times g$ a 4°C y luego se transfirió el sobrenadante transparente a un nuevo tubo de 1,50 mL libre de ARNasas. Este paso es esencial para aquellas muestras que presentan alto contenido en grasa, como es el caso de las muestras analizadas en este estudio (170).
6. Se incorporaron 200 μL de cloroformo (cat. N°: 1.02045, Sigma-Aldrich, St. Louis, USA) y se agitó en el vórtex durante 25 segundos. Pasado este tiempo, se incubó a temperatura ambiente durante 3 minutos.
7. Se centrifugó la muestra a $12.000 \times g$ durante 15 minutos a 4°C . Al finalizar la centrifugación se observó una fase roja inferior de fenol-cloroformo, una interfase de color blanquecino y una fase acuosa superior incolora que contenía el ARN.
8. Se transfirieron 550 μL de la fase acuosa que contenía el ARN a un tubo nuevo libre de ARNasas y se incorporó un volumen igual de etanol al 70% (cat. N°: ET00110500, Scharlab, Barcelona, España). Para mezclar bien el etanol con la fase acuosa, se agitó en el vórtex durante 10 segundos.
9. Posteriormente, se transfirieron $\leq 700 \mu\text{L}$ de la muestra a una columna de centrifugado con su correspondiente tubo colector suministrados por el kit PureLink®.
10. Se centrifugó a $12.000 \times g$ durante 25 segundos a temperatura ambiente. El líquido presente en el tubo colector se desechó y se volvió a insertar la columna en el mismo tubo colector.
11. Se repitieron los pasos 9 y 10 hasta procesar toda la muestra.

12. Se añadieron 350 μL de tampón de lavado I suministrado por el kit PureLink® a la columna de centrifugado. Se centrifugó a 12.000 $\times g$ durante 25 segundos a temperatura ambiente. El líquido y el tubo colector se desecharon y se insertó a la columna un nuevo tubo colector.
13. Se incorporaron 80 μL de mezcla PureLink® DNasa directamente sobre la superficie de la membrana de la columna de centrifugado y se incubó a temperatura ambiente durante 15 minutos. Esta etapa permitió la degradación del ADN presente. La mezcla PureLink® DNasa consistió en: 8 μL de 10X DNase I Reaction Buffer, 10 μL de Resuspended DNase ($\sim 3\text{U}/\mu\text{L}$) y 62 μL de agua libre de ARNasas.
14. Al finalizar el tiempo de incubación, se añadieron 350 μL de tampón de lavado I PureLink® a la columna de centrifugado. Se centrifugó a 12.000 $\times g$ durante 25 segundos a temperatura ambiente. El líquido y el tubo colector se desecharon y se insertó la columna un nuevo tubo colector.
15. Se añadieron 500 μL de tampón de lavado II suministrado por el kit PureLink® a la columna de centrifugado PureLink®.
16. Se centrifugó a 12.000 $\times g$ durante 25 segundos a temperatura ambiente y se desechó el líquido presente en el tubo colector. La columna se volvió a insertar en el mismo tubo colector.
17. Se repitieron los pasos 15 y 16.
18. El tampón de lavado II PureLink® contiene etanol, por lo que se centrifugó a 12.000 $\times g$ durante 1 minuto a temperatura ambiente con el fin de eliminar los restos de etanol. Se desechó el tubo colector y se insertó la columna de centrifugado en un tubo de recuperación.
19. Se incorporaron 16 μL de agua libre de ARNasas en el centro de la columna de centrifugado PureLink® y se incubó a temperatura ambiente durante 1 minuto.
20. Se centrifugó la columna PureLink® a $\geq 12.000 \times g$ durante 1 minuto y a temperatura ambiente.
21. Los pasos 19 y 20 se repitieron una vez más.
22. Finalmente, se realizó una dilución 1/10 para poder llevar a cabo la cuantificación fluorimétrica. Esta dilución se realizó incorporando 1 μL del ARN total extraído en 9 μL de agua libre de ARNasas. Además, se traspasó 3 μL del ARN total extraído a otro tubo para determinar su calidad mediante una cuantificación espectrofotométrica (1 μL) y para migrarlo en un gel de agarosa (2 μL). El resto de ARN total se almacenó a $-150\text{ }^\circ\text{C}$.

3.2.2. Cuantificación y análisis de calidad del ARN total

Para poder valorar la cantidad de ARN total extraído se llevó a cabo una cuantificación fluorimétrica mediante Qubit® RNA HS Assay Kit (cat. N°: Q32852, Invitrogen, Carlsbad, USA). Para ello, se preparó la solución de trabajo consistente en una dilución 1:200 del Qubit® RNA HS reagent en Qubit® RNA HS buffer. Posteriormente se llevó a cabo una recta patrón utilizando 10 µL de cada muestra estándar suministrada por el kit y 190 µL de la solución de trabajo. La muestra estándar N° 1 no contenía ARN, mientras que la muestra estándar N° 2 presentaba una concentración de ARN de 10 ng/µL. La cuantificación de las muestras de ARN total se realizó añadiendo 1 µL de la dilución 1/10 en 199 µL de la solución de trabajo. La cuantificación se llevó a cabo en el Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, USA) siguiendo las instrucciones del fabricante.

Para el análisis de calidad del ARN total, se realizó una cuantificación espectrofotométrica mediante el Scientific NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, Waltham, USA). Este procedimiento, además de cuantificar la cantidad de ARN presente en una muestra, se utiliza para detectar la presencia de algunos contaminantes procedentes del proceso de extracción del ARN, como son, restos de alcoholes y fenoles, proteínas o sales disueltas en el medio. El ARN presenta un pico de absorción a 260 nm (A260), mientras que contaminantes como las proteínas o las sales presentan picos de absorción a 280 nm (A280) y 230 nm (A230), respectivamente. La cercanía de los máximos de absorción del ARN, las sales y las proteínas hace que en algunas muestras haya un solapamiento parcial, por lo que es necesario una estimación de la relación A260/A280 y A260/A230 para obtener una mejor consideración de la cantidad y calidad de ARN. Valores en la relación A260/A280 de entre 1,8 y 2 indican grados de eliminación de proteína óptimos, mientras que, para las sales, la relación A260/A230 tiene sus valores óptimos entre 1,8 y 2,2. En ambos casos, valores menores de 1,8 indican muestras con un mayor nivel de proteínas y sales y menor pureza.

Adicionalmente, se procedió a valorar la degradación del ARN total mediante migración en un gel de agarosa (171). Para ello se preparó un gel de agarosa al 1,50% (cat. N°: A2114,0500, PanReac AppliChem ITW Reagents, Germany) en solución TBE 1X. Al gel se le incorporó 0,003% de GelRed™ (cat. N°: 41003, Biotium, California, USA) que en contacto con el ARN y, al ser irradiado con luz UV de 254 nm, emite fluorescencia indicando la presencia de ARN. En cada pocillo se depositaron 2 µL de ARN total junto con 1,50 µL de solución de carga y 1,50 µL de un

marcador de peso molecular de ADN de 100 pb (cat. N°: 31.006, Biotools, Madrid, España) utilizado como control de la migración. Posteriormente, se realizó la migración electroforética en una cuba Wide-Sub Cell GT con una fuente de alimentación PowerPac Basic (cat. N°: 1640301, Bio-Rad, California, USA), durante 30 minutos y a 100 V constante. Al finalizar el tiempo de migración, el gel fue visualizado en un UV transiluminador UVI-DOC v15.01 (UVItec, Cambridge, UK).

Soluciones para la preparación de geles de agarosa

Solución de carga (azul de carga):

- EDTA 0,50 M pH 8 (cat. N°: E5134, Sigma-Aldrich, St. Louis, USA),
- glicerol 80% v/v,
- azul de bromofenol 25% w/v (cat. N°: B-6896, Sigma-Aldrich, St. Louis, USA),
- xilencianol 25% w/v.

Solución TBE 10X:

- 108 g de Tris (cat. N°: 75825, Affymetrix, High Wycombe, UK),
- 55 g de ácido bórico (cat. N°: 131015, PanReac AppliChem ITW Reagents, Germany),
- 40 mL de EDTA a 0,50 pH 8,
- Agua milliQ (Merck Millipore, Germany) hasta el volumen de 1 L.
- Antes de ser utilizado el TBE debe ser autoclavado. El TBE 1X se prepara por dilución de 100 mL de TBE 10X con 900 mL de agua milliQ.

3.3. ANÁLISIS DE LA METILACIÓN m⁶A

3.3.1. Determinación del porcentaje de metilación m⁶A global

La determinación del porcentaje m⁶A se llevó a cabo con el kit EpiQuik m⁶A RNA Methylation Quantification Kit (Colorimetric) (cat. N°: P-9005-96, Epigentek, NY, USA), siguiendo las condiciones del fabricante. En este ensayo, el ARN total se une al fondo de los pocillos de la placa mediante el uso de una solución de alta unión. El m⁶A es detectado mediante anticuerpos de captura y detección, y la señal que emite es amplificada y cuantificada colorimétricamente en un espectrofotómetro de placas. Finalmente, se realizan los cálculos de la cantidad y

porcentaje de m⁶A teniendo en cuenta que la cantidad de m⁶A es proporcional a la intensidad de absorbancia (OD).

Para llevar a cabo la determinación del porcentaje de metilación m⁶A global, se seleccionaron 20 muestras tumorales y sus correspondientes tejidos mamarios sanos para ser analizados por duplicado. Únicamente evaluamos este número de muestras ya que es el máximo que se puede analizar simultáneamente en la misma placa en la que se obtiene la curva estándar. Las muestras seleccionadas pertenecían a los distintos subtipos moleculares de cáncer de mama, siendo por cada subtipo cuatro muestras tumorales y cuatro muestras correspondientes al tejido sano. Por cada muestra se utilizaron 200 ng de ARN total. Además de las muestras de tejido mamario, también se analizaron por duplicado la muestra de control negativo (un ARN que no contiene m⁶A) y las diluciones a distintas concentraciones del control positivo (oligos de m⁶A normalizados para tener el 100% de m⁶A). Ambos controles, negativo y positivo, fueron suministrados por el kit. La señal se cuantificó espectrofotométricamente mediante la lectura de la absorbancia a una longitud de onda de 450 nm en un GloMax™-Multi+ Detection System (cat. N°: TM297, Promega, Madison, USA). El porcentaje de metilación m⁶A del ARN se calculó de manera absoluta. Para ello, se generó una curva estándar a partir de controles positivos suministrados con el kit. La cantidad final de los controles positivos utilizados fue de 0,01, 0,02, 0,04, 0,10, 0,20, 0,40 y 1 ng por pocillo. Se representó gráficamente el valor promedio de OD de cada control positivo frente a la cantidad final de dichos controles. A continuación, se determinó la pendiente (OD/ng) de la curva estándar mediante regresión lineal utilizando Microsoft Excel (2016). Para un cálculo óptimo de la pendiente, se utilizó la parte más lineal de la curva, la cual incluye al menos 4 puntos. Posteriormente, se calcularon la cantidad y el porcentaje de m⁶A para cada muestra mediante las fórmulas 3.3.1 y 3.3.2:

$$\text{Cantidad de m6A (ng)} = \frac{\text{OD de la muestra} - \text{OD del control negativo}}{\text{pendiente de la recta}} \quad (3.3.1)$$

$$\text{Porcentaje de m6A (\%)} = \frac{\text{cantidad de m6A (ng)}}{200 \text{ ng}} \times 100 \quad (3.3.2)$$

Donde los 200 ng se corresponden con la cantidad inicial de ARN total de las muestras.

Una vez obtenido el porcentaje de metilación m⁶A para cada muestra, se procedió al análisis estadístico de los datos mediante el programa IBM SPSS Statistics 25 para Windows (SPSS Inc., IBM Corporation, NY, USA). Para determinar la existencia de diferencias significativas entre el porcentaje de metilación m⁶A de las muestras de cáncer de mama y sus correspondientes tejidos sanos, se realizó la prueba de la U de Mann-Whitney. Por otra parte, para determinar las diferencias entre los distintos subtipos tumorales, se llevó a cabo la prueba estadística H de Kruskal-Wallis. Los valores de *p* inferiores a 0,05 fueron considerados estadísticamente significativos.

3.3.2. Análisis de expresión génica de las metiltransferasas y desmetilasas

El análisis de expresión génica mediante RT-qPCR se llevó a cabo con el fin de determinar la expresión de los genes que codifican para las metiltransferasas (*METTL3*, *METTL14* y *WTAP*) y desmetilasas (*FTO* y *ALKBH5*). La expresión de dichos genes se normalizó mediante la amplificación de los genes *housekeeping* gliceraldehído-3-fosfato deshidrogenasa (*GADPH*), β -actina (*ACTB*) y β -glucuronidasa (*GUSB*).

3.3.2.1. Generación del ADNc a partir del ARN total

La generación del ADNc de todas las muestras analizadas en este estudio se realizó mediante el kit SuperScript™ VILO™ cDNA Synthesis Kit (cat. N°: 11754050, Invitrogen, Carlsbad, USA). Para ello, se mezclaron 4 μ L de 5X VILO™ Reaction Mix, 2 μ L de 10X SuperScript™ Enzyme Mix, 250 ng de ARN total y agua libre de ARNasas (cat. N°: 10977049, Invitrogen, Carlsbad, USA) hasta llegar a un volumen final de 20 μ L. La reacción se llevó a cabo en un termociclador C1000 Touch™ Thermal Cycler (cat. N°: 1851148, Bio-Rad, California, USA) bajo las siguientes condiciones: 10 minutos a 25 °C, 60 minutos a 42 °C, 5 minutos a 85 °C y un ciclo final de 12 °C.

3.3.2.2. Cuantificación del ADNc

Para poder valorar la cantidad de ADNc generado, se realizó una cuantificación fluorimétrica con el Qubit® 1X dsDNA HS Assay Kit (cat. N°: Q33230, Invitrogen, Carlsbad, USA). Para ello, primero se llevó a cabo una recta patrón, utilizando 10 μ L de cada muestra estándar suministrada por el kit y 190 μ L de la solución de trabajo. La muestra estándar N° 1 no contenía

ADN, mientras que la muestra estándar N° 2 presentaba una concentración de ADN de doble hebra de 10 ng/μL. La cuantificación de las muestras de ADNc se realizó añadiendo 1 μL a 199 μL de la solución de trabajo. La cuantificación se llevó a cabo en un Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, USA), siguiendo las instrucciones del fabricante. Posteriormente, el ADNc generado fue diluido a una concentración de 2,50 ng/μL.

3.3.2.3. Diseño de los cebadores

Las parejas de cebadores para los genes *METLL3*, *WTAP*, *FTO* y *GADPH* fueron diseñadas utilizando el programa PerlPrimer v1.1.21 (172). Debido al proceso de empalme alternativo, cada gen puede dar origen a diversos transcritos que se podrían traducir o no a proteínas. Por tal motivo, el diseño de los cebadores fue realizado mediante la selección de una región presente en todos los transcritos. En caso de no existir esta región común a todos, se seleccionaron regiones que se encontraban en la mayoría de los transcritos. Por otra parte, las parejas de cebadores de los genes *METTL14*, *ALKBH5*, *GUSB* y *ACTB* fueron obtenidas a partir de la bibliografía (173,174). Todas las parejas de cebadores fueron verificadas mediante el software Primer-BLAST del NCBI (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) (175) y sintetizados por Metabion (Planegg, Germany) (Tabla 3.3.1).

Tabla 3.3.1. Secuencia, exones donde hibridan y tamaño del amplicón de los cebadores utilizados en este estudio.

Gen	Secuencia 5'-3' de los cebadores	Exón donde hibridan	Tamaño del amplicón (pb)	Ref.
<i>METTL3</i>	F: GAGGAGTGCATGAAAGCCAG R: GGCCTCAGAATCCATGCAAG	4 5	180	Presente estudio
<i>METTL14</i>	F: GTTGAACATGGATAGCCGC R: CAATGCTGTCGGCACTTCA	1 2	102	(173)
<i>WTAP</i>	F: GCCTGATAGCCAAACAGGGAAA R: TCATTCAGTTCATCCTGACTGC	6-7 7-8	174	Presente estudio
<i>FTO</i>	F: GCCAGAACCTGAGGAGAGAAT R: AGGTCAAACGGCAGAGGC	8 9	133	Presente estudio
<i>ALKBH5</i>	F: ATGCACCCCGTTGGAAAC R: GACTTGCGCCAGTAGTTCTCA	3 4	250	(173)
<i>GUSB</i>	F: GAAAATACGTGGTTGGAGAGCTCATT R: CCGAGTGAAGATCCCTTTTCA	9 10	101	(174)
<i>GADPH</i>	F: CACTGGCGTCTTCACCAC R: GATGGCATGGACTGTGGT	5 8	250	Presente estudio
<i>ACTB</i>	F: CCTGACTGACTACCTCATGAAG R: GACGTAGCACAGCTTCTCCTTA	4 4	101	(173)

3.3.2.4. Análisis de eficiencia

Para evaluar la eficiencia de las parejas de cebadores seleccionadas para cada gen, se llevó a cabo el procedimiento descrito por King (2.007) (176). Para ello:

- I. Se realizó una amplificación por PCR convencional de cada fragmento de los distintos genes utilizando el ADNc generado a partir del ARN total de muestras de tejido sano. La reacción de PCR consistió en 17,10 μL de agua milliQ, 0,50 μL de dNTP (2,50 mM) (cat. N°: BIO-39028, Bioline, Barcelona, España), 0,75 μL de MgCl_2 (50 mM) (cat. N°: BIO-21040, Bioline, Barcelona, España), 2,50 μL de tampón 10X (cat. N°: BIO-21040, Bioline, Barcelona, España), 0,25 μL de cada cebador para cada gen (10 μM), 1 μL de albúmina de suero bovino (cat. N°: 10711454001, Sigma-Aldrich, St. Louis, USA), 0,15 μL de ADN polimerasa BIOTAQ™ (5 U/ μl) (cat. N°: BIO-21040, Bioline, Barcelona, España) y 2,50 μl de ADNc (2,50 ng/ μl). La PCR se llevó a cabo en un termociclador C1000 Touch™ Thermal Cycler. Las condiciones fueron: desnaturalización inicial a 98 °C durante 2 minutos, seguida de 35 ciclos de 10 segundos a 98 °C, 30 segundos a 61 °C y 1 minuto a 72 °C, y una extensión final de 5 minutos a 72 °C.
- II. Todo el volumen de reacción del paso 1 se migró en un gel de agarosa al 1,50%. La banda obtenida, tras verificar que se ajustaba al tamaño esperado, fue purificada mediante el kit QIAEX II Gel Extraction (cat. N°: 20021, Qiagen, Hilden, Germany) siguiendo las instrucciones del fabricante.
- III. El producto obtenido tras la purificación de las bandas del gel se cuantificó fluorimétricamente mediante el kit Qubit® 1X dsDNA HS Assay Kit, siguiendo el procedimiento descrito en el apartado 3.3.2.2.
- IV. Una vez obtenida la concentración del producto de expresión génica, se determinó el número de copias utilizando la fórmula 3.3.3:

$$N^{\circ} \text{ de copias (copias}/\mu\text{L}) = \frac{6,022 \times 10^{23} \left(\frac{\text{copias}}{\text{mol}}\right) \times \text{concentración} \left(\frac{\text{g}}{\text{mol}}\right)}{\text{peso molecular} \left(\frac{\text{g}}{\text{mol}}\right)} \quad (3.3.3)$$

Donde, la concentración (g/ μL) se determinó en el paso III y el peso molecular (g/mol) se calculó a través del software <http://www.lab314.com/genmol/dsdna.htm>.

- V. Finalmente, se realizaron diluciones seriadas con distintos números de copias y se llevó a cabo la RT-qPCR. Cada dilución se analizó por triplicado. La eficiencia, así como el valor del coeficiente de correlación (R^2), la pendiente y el punto de intersección con el eje Y se obtuvieron mediante el programa CFX Manager™ 2.0. (cat. N°: 1845000, Bio-Rad, California, USA).

3.3.2.5. RT-qPCR

La amplificación mediante RT-qPCR se realizó en un termociclador C1000™ acoplado al sistema de detección a tiempo real CFX96™ (cat. N°: 1855196, Bio-Rad, California, USA). Tras la reacción de RT-qPCR se llevó a cabo la técnica de High Resolution Melting (HRM) con el fin de verificar la presencia o ausencia de contaminación mediante el análisis de los picos de fusión.

La reacción de RT-qPCR se llevó a cabo en un volumen final de 5 μ l con los siguientes reactivos: 2,50 μ L de SsoFastST EvaGreen® Supermix (cat. N°: 1725200, Bio-Rad, California, USA), 0,50 μ L de cada cebador (1 μ M), 0,50 μ L de agua milli-Q y 1 μ L de ADNc (2,50 ng/ μ L). Las condiciones fueron: desnaturalización inicial a 98 °C durante 2 minutos, seguida de 35 ciclos a 98 °C durante 10 segundos y a 61 °C durante 30 segundos. Las curvas de fusión se obtuvieron de 65 °C a 95 °C con un incremento de 0,20 °C/segundo. La RT-qPCR de cada gen se realizó por triplicado en muestras de ADNc de cáncer de mama, así como en sus correspondientes tejidos sanos. El análisis de los resultados obtenidos y la determinación del ciclo umbral (Cq) se realizaron con el programa CFX Manager™ 3.1.

3.3.2.6. Análisis de expresión génica

El análisis de expresión génica se calculó mediante el Cq promedio de cada muestra amplificada por triplicado. Este valor fue el utilizado para estimar los niveles de expresión génica relativa para cada muestra mediante el método de $2^{-\Delta\Delta Cq}$ (177). Para ello, primero se calculó el ΔCq , que es igual a la resta del Cq promedio de los genes *housekeeping* menos el Cq promedio del gen de interés. Posteriormente, se calculó el $\Delta\Delta Cq$, que es igual a la resta del ΔCq obtenido para cada muestra menos el ΔCq del calibrador. En nuestro caso, utilizamos como calibrador el ΔCq promedio de las muestras de tejido mamario sano. Finalmente, se llevó a cabo la potencia de dos elevado a $-\Delta\Delta Cq$.

3.3.2.7. Análisis estadístico

Se calcularon el promedio de la expresión génica y la desviación estándar para cada muestra. Para determinar si había diferencias en la expresión de cada gen entre las muestras tumorales y sus correspondientes tejidos sanos, se llevó a cabo la prueba estadística U de Mann-Whitney. La diferencia entre los distintos subtipos tumorales se realizó con la prueba H de Kruskal-Wallis. También se evaluó la correlación entre los distintos genes mediante el coeficiente de correlación de Pearson. Todos los análisis estadísticos se llevaron a cabo con el programa IBM SPSS Statistics 25 para Windows (SPSS Inc., IBM Corporation, NY, USA). Los valores de p inferiores a 0,05 se consideraron estadísticamente significativos.

3.4. IDENTIFICACIÓN *IN SILICO* Y ANÁLISIS DE EXPRESIÓN *IN VITRO* DE m⁶A-SNPs

Recientemente, se ha visto que los m⁶A-SNPs podrían afectar la expresión de los genes según el tejido y, en consecuencia, influir en el desarrollo y progresión de distintas enfermedades. Con esta premisa, se ha llevado a cabo una búsqueda de m⁶A-SNPs asociados al cáncer de mama y se ha evaluado *in silico* e *in vitro* los posibles efectos de esta marca epigenética en la expresión génica en esta enfermedad.

3.4.1. Determinación de m⁶A-SNPs relacionados con el cáncer de mama

Para poder determinar el efecto de los m⁶A-SNPs en el cáncer de mama, primeramente, se identificaron los SNPs relacionados con dicha enfermedad. Para ello, se buscaron estos SNPs en las siguientes bases de datos: NCI Genomic Data Commons (GDC, <https://portal.gdc.cancer.gov/>) (178), GWAS catalog (<https://www.ebi.ac.uk/gwas/>) (179) y SNPedia (<https://www.snpedia.com/>) (180). Para ampliar el estudio, también se buscaron los SNPs presentes en los miARN relacionados con cáncer de mama en las siguientes bases de datos: miRdSNP (<http://mirdsnp.ccr.buffalo.edu/index.php>) (181), miRNASNP (<http://bioinfo.life.hust.edu.cn/miRNASNP/#!/>) (182), MSDD (<http://bio-bigdata.hrbmu.edu.cn/msdd/home.jsp>) (183) y SomamiR (<http://compbio.uthsc.edu/SomamiR/>) (184). Posteriormente, se descargó una lista de m⁶A-SNPs y sus datos asociados presentes en la base de datos m⁶Avar (<http://m6avar.renlab.org/help.html>) (145). Esta base de datos contiene m⁶A-SNPs humanos

con diferentes niveles de confianza (alto, medio y bajo). Finalmente, se comparó la lista de los SNPs asociados al cáncer de mama con la de los m⁶A-SNPs para identificar cuáles aparecen en ambas listas.

Los datos sobre las frecuencias alélicas y genotípicas para los m⁶A-SNPs identificados se obtuvieron de Ensembl Genome Browser (<https://www.ensembl.org/index.html>) (185).

3.4.2. Análisis de loci de rasgos cuantitativos de expresión de los m⁶A-SNPs

Para determinar si los m⁶A-SNPs cambian la expresión de los genes, se llevó a cabo una búsqueda de análisis *cis*-eQTL utilizando los datos presentes en el Portal GTEx (<https://gtexportal.org/home/>) (186) y en el navegador HaploReg (<http://archive.broadinstitute.org/mammals/haploreg/haploreg.php>) (187). También, se buscaron las funciones potenciales en la regulación de la expresión génica, como la alteración de la unión a proteínas o el motivo de regulación en el navegador HaploReg.

3.4.3. Análisis de expresión génica *in silico*

Una vez determinados los m⁶A-SNPs con efecto *cis*-eQTL, se evaluó si sus genes presentaban una expresión diferencial en el cáncer de mama. Con este fin, en el caso de los m⁶A-SNPs presentes en el ARNm, se descargaron dos conjuntos de datos que contenían los valores de expresión génica normalizados para el tejido mamario de la base de datos GEO (GSE5764 y GSE21422) (<http://www.ncbi.nlm.nih.gov/geo>) (188). El conjunto GSE5764 contenía los datos de los niveles de expresión génica de 10 individuos con carcinoma de mama (cinco con carcinoma de mama ductal invasivo y cinco con carcinoma de mama lobulillar invasivo) y 20 correspondientes al tejido normal (10 al tejido ductal normal y 10 al tejido lobulillar normal). Por otra parte, el conjunto GSE21422 contenía datos de cinco muestras de tejido sano, nueve carcinomas ductales *in situ* y cinco carcinomas ductales invasivos. Las diferencias en los niveles de expresión génica se llevaron a cabo mediante la prueba U de Mann-Whitney a través del programa IBM SPSS Statistics 25.0 (SPSS Inc., IBM Corporation, Nueva York, EE. UU.). Se consideraron como genes de expresión diferencial si presentaban valores de $p < 0,05$ en al menos uno de los dos estudios.

Por otra parte, en cuanto a los m⁶A-SNPs presentes en los miARN, los datos sobre el número de muestras, la media de la expresión génica y los valores de *p* se obtuvieron de la base de datos ENCORI para ARN Interactomes (<http://starbase.sysu.edu.cn/panMirDiffExp.php>) (189).

3.4.4. Validación de las posiciones m⁶A-SNPs

3.4.4.1. Muestras a analizar y diseño de las parejas de cebadores

Para validar los m⁶A-SNPs que mostraron un efecto *cis*-eQTLs y diferencias en la expresión génica en el análisis *in silico*, se procedió al análisis de las muestras de las pacientes con cáncer de mama de este estudio. Para ello, se analizó el ADNc de las 35 muestras tumorales descritas en el apartado 3.1 con el fin de determinar sus genotipos para cada m⁶A-SNPs y poder compararlo con la población control. Por otra parte, se descargaron los datos de poblaciones Ibéricas en España (IBS) para los m⁶A-SNPs a partir de la base de datos *1000 Genomes* (<https://www.internationalgenome.org/home>) (190) con el fin de poder utilizarla como población control. Para este análisis únicamente se seleccionaron los datos procedentes de mujeres.

Los cebadores para los m⁶A-SNPs identificados en el estudio *in silico* fueron diseñados en este estudio utilizando el programa PerlPrimer v1.1.21 (172). Las secuencias de los cebadores, así como el tamaño del amplificado, se muestran en la Tabla 3.4.1. Todas las parejas de cebadores fueron verificadas mediante el software Primer-BLAST del NCBI (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) (175) y sintetizados por Metabion (Planegg, Germany). Las parejas de cebadores fueron amplificadas a diferentes temperaturas de hibridación para determinar la temperatura óptima a la que se obtenía la mayor cantidad de producto amplificado.

Tabla 3.4.1. Información acerca de los cebadores utilizados en este estudio.

m ⁶ A-SNP	Gen	Secuencia 5'-3' de los cebadores	T de hibridación (°C)	Amplión (pb)
rs1801270	<i>CDKN1A</i>	F: GGGATTTCTTCTGTTCAGGCG* R: TCGAAGTTCCATCGCTCACG	57	178
rs76563149	<i>ZNF354A</i>	F: GGGTCGCGAGGCTGCATC R: AACGTCAGTGACACCTGGG*	61	181
rs11614913	<i>MIR196A2</i>	F: GCTCAGCTGATCTGTGGCTTA* R: CCTCGACGAAAACCGACTGA	61	105

*Cebadores utilizados para la secuenciación Sanger.

3.4.4.2. *Secuenciación Sanger*

Con el fin de determinar el genotipo de las muestras de cáncer de mama de cada m⁶A-SNPs, se procedió a secuenciarlas a través del siguiente procedimiento.

1. *Reacción de PCR de amplificación.* La amplificación PCR de cada m⁶A-SNP se llevó a cabo en un termociclador C1000™ acoplado al sistema de detección a tiempo real CFX96™ (cat. N°: 1855196, Bio-Rad, California, USA). La reacción de amplificación contenía los siguientes reactivos: 2,50 µL de SsoFastST EvaGreen® Supermix (cat. N°: 1725200, Bio-Rad, California, USA), 0,50 µL de cada cebador a una concentración de 1 µM, 0,50 µL de agua milli-Q y 1 µL de ADNc (2,50 ng/µL). Las condiciones de amplificación utilizadas fueron: desnaturalización inicial a 98 °C durante 2 minutos seguida de 35 ciclos a 98 °C durante 10 segundos y a la temperatura óptima de cada pareja de cebadores (Tabla 3.4.1) durante 30 segundos. La amplificación de las muestras fue verificada mediante el programa CFX Manager™ 3.1 software (cat. N°: 1845000, Bio-Rad, California, USA).
2. *Purificación de los productos amplificados.* La purificación de los productos de amplificación se realizó mediante el reactivo ExoSAP-IT™ Express (cat. N°: 75001, Applied Biosystems, Waltham, USA), que contiene los enzimas: Exonucleasa I (Exo I) y Shrimp alkaline phosphatase (SAP). La limpieza con Exo-SAP-IT™ se llevó a cabo en el termociclador C1000 Touch™ Thermal Cycler (cat. N°: 1851148, Bio-Rad, California, USA) mediante la incorporación de 1 µL del reactivo en 2,50 µL del producto amplificado. Las muestras se incubaron a 37 °C durante 4 minutos para que las enzimas actuaran y a 80 °C durante 1 minuto para desactivarlas.
3. *Reacción de PCR de secuenciación.* A diferencia de la reacción de PCR de amplificación, la reacción de PCR de secuenciación incorpora dNTPs y ddNTPs a la reacción. Los ddNTPs se caracterizan por estar marcados con distintos fluoróforos, uno para cada nucleótido. Dado que los ddNTPs no poseen un grupo OH en la posición 3' de la ribosa, detienen la elongación del ADN. La incorporación de un ddNTP es aleatoria, por lo que se obtienen amplificados de todos los tamaños posibles lo que permite obtener la secuencia completa del fragmento de ADN previamente amplificado. La reacción de secuenciación se llevó a cabo en un termociclador C1000 Touch™ Thermal Cycler (cat. N°: 1851148, Bio-Rad, California, USA) mezclando 0,70 µL de BigDye® Terminator™ v.1.1 Cycle Sequencing Kit (cat. N°: 4337450, Applied Biosystems, Waltham, USA), 0,25 µL del cebador más alejado del m⁶A-SNP de interés (10 mM) (Tabla 3.4.1), 5,55 µL de agua milli-Q y 3,5 µL del amplificado de PCR limpio.

El protocolo utilizado fue: un ciclo inicial a 96 °C durante 1 minuto, seguido de 25 ciclos de 10 segundos a 96 °C, 5 segundos a 55 °C y 1 minuto y 15 segundos a 60 °C.

4. *Purificación de los productos de secuenciación.* En esta purificación se utilizó el Big Dye® Xterminator™ purification Kit (cat. N°: 75001, Applied Biosystems, Walthman, USA) para eliminar los productos sobrantes de la reacción de PCR de secuenciación tales como sales, dNTPs y ddNTPs y así evitar posibles interferencias en los resultados de la secuenciación. Para ello, el BigDye® Xterminator™ purificación Kit se incorporó directamente al producto de PCR de secuenciación, añadiendo 22,50 µL de la solución SAM™ y 5,50 µL de la solución Xterminator™. Posteriormente, se sometió a agitación a 2.000 rpm en un agitador Eppendorf™ MixMate™ (cat. N°: 5353000014, Eppendorf, Madrid, España), durante 50 minutos y se centrifugó durante 4 minutos a 1.000 x *g*. El sobrenadante, que contiene los productos amplificados, fue recogido para su posterior análisis.
5. *Electroforesis capilar.* Los productos amplificados y purificados se migraron mediante electroforesis capilar en un secuenciador automático ABI 3130 (cat. N°: 3130-01, Applied Biosystems, Walthman, USA) con un capilar de 36 cm y el polímero POP7™ (cat. N°: 4363785, Applied Biosystems, Walthman, USA). Para realizar la migración electroforética, se añadieron 5 µL del producto de PCR de secuenciación purificado y 5 µL de HI-DI™ formamide (cat. N°: 4363785, Applied Biosystems, Walthman, USA).
6. *Análisis de los resultados.* El análisis inicial de las secuencias obtenidas, así como, la asignación de las bases nitrogenadas, se llevaron a cabo con el programa informático Sequencing Analysis software v.5.2 (Applied Biosystems, Walthman, USA). Posteriormente, el genotipo de cada muestra fue determinado mediante el programa ChromasPro v.2.6.5 (Technelysium Pty Ltd, Brisbane, Australia).

3.4.4.3. *Análisis de expresión génica*

Para aquellos m⁶A-SNPs que presentaron diferencias en cuanto las frecuencias genotípicas y/o alélicas, se procedió a realizar un análisis de expresión en las muestras tumorales y sus correspondientes tejidos sanos. Para ello, primeramente, se determinó la eficiencia de los cebadores siguiendo el procedimiento descrito en el apartado 3.3.2.4. Posteriormente se llevó a cabo el análisis de expresión tal y como se describe en los apartados 3.3.2.5 y 3.3.2.6 utilizando como genes *housekeeping* a *GADPH*, *ACTB* y *GUSB*.

3.4.4.4. *Análisis estadísticos*

Las frecuencias alélicas y genotípicas de cada m⁶A-SNP fueron calculadas por contaje directo. Las diferencias entre frecuencias del grupo de pacientes con cáncer de mama y del grupo de población general se analizaron mediante la prueba de chi-cuadrado (χ^2) de Pearson. Por otra parte, las diferencias en cuanto a la expresión génica entre las muestras tumorales y sus correspondientes tejidos sanos fueron analizados mediante la prueba estadística de U de Mann-Whitney. La diferencia entre los diferentes subtipos tumorales se llevó a cabo mediante la prueba estadística H de Kruskal-Wallis. Para todas las pruebas se utilizó el software IBM SPSS Statistics 25.0 (SPSS Inc., IBM Corporation, NY, EE. UU.). Los valores de $p < 0,05$ se consideraron estadísticamente significativos.

3.5. DETERMINACIÓN DE POSICIONES m⁶A METILADAS MEDIANTE LA TÉCNICA MeRIP

La identificación de posiciones m⁶A metiladas presentes en las muestras de tejidos tumorales de cada subtipo molecular se llevó a cabo mediante la técnica MeRIP desarrollada por Dominissini *et al.* (2.013) (85) con algunas modificaciones propuestas por Zheng *et al.* (2.018) (89). Resumidamente, esta técnica consta de las siguientes etapas: 1) fragmentación del ARN total, 2) inmunoprecipitación del ARN total utilizando anticuerpos anti-m⁶A, 3) cuantificación del m⁶A-ARN inmunoprecipitado, 4) generación, cuantificación y evaluación de la librería, 5) secuenciación de la librería mediante *Next Generations Sequencing* (NGS) y 6) análisis de los datos obtenidos tras la secuenciación.

La identificación de las posiciones m⁶A metiladas fue efectuada en 20 muestras tumorales, cuatro por cada subtipo molecular.

3.5.1. *Fragmentación del ARN total*

La fragmentación del ARN total se llevó a cabo de manera química con el kit RNA Fragmentation Reagents (cat. N^o: AM8740, Invitrogen, Carlsbad, USA), para obtener fragmentos de tamaños de 100-250 pb. Para ello, se mezclaron en hielo, 18 μ L que contenían 10 μ g de ARN total con 2 μ L de Ambion[®] 10X RNA Fragmentation Reagent y se incubaron a 70 °C durante 4 minutos en un termoagitador (cat. N^o: PHMT-PSC18, Grant Instruments[™], Cambridge, UK). Al finalizar el

tiempo de incubación, se incorporaron rápidamente 2 μL de Stop solution para detener la reacción. También se llevó a cabo la fragmentación de un control negativo, es decir, la fragmentación de una muestra de agua libre de ARNasas y sin ARN. El tamaño de los fragmentos se confirmó mediante la migración de 0,75 μL de cada ARN fragmentado en un gel de agarosa al 1,50 % a voltaje constante de 100 V durante 30 minutos.

No todo el ARN fragmentado continuó hacia el proceso de inmunoprecipitación, sino que 1,50 μL del ARN fragmentado (que en adelante llamaremos INPUT) se diluyó en 13,50 μL de agua libre de ARNasas y se almacenó para su posterior cuantificación y generación de librería.

3.5.2. Inmunoprecipitación

Finalizada la fragmentación, se continuó con la etapa de inmunoprecipitación. El protocolo para llevar a cabo la inmunoprecipitación fue el siguiente:

1. Se mezclaron los siguientes reactivos: 470 μL de tampón de baja unión, 10 μL del anticuerpo anti-m⁶A (cat. N^o: ab151230, Abcam, Cambridge, UK), 2 μL del inhibidor de ARNasas (cat. N^o: N8080119, Applied Biosystems, Waltham, USA), y 19,75 μL del ARN fragmentado/control negativo. Se mantuvo en un agitador rotatorio (cat. N^o: LJD001, Nahita, Navarra, España) toda la noche a 4 °C. De esta manera, los fragmentos de ARN con posiciones m⁶A se unen al anticuerpo anti-m⁶A.
2. Al día siguiente, mientras el paso 1 está en curso, se dejaron a temperatura ambiente durante 20 minutos microesferas magnéticas de proteína A/G Pierce (cat. N^o: 88802, Thermo Scientific, Waltham, USA). Posteriormente, se transfirieron 25 μL de las microesferas magnéticas por muestra a un tubo *LoBind* limpio de 1,50 μL (cat. N^o: 0030108035, Eppendorf, Madrid, España) y se lavaron dos veces con 1 mL de tampón de baja unión. El lavado se realizó colocando el tubo que contenía el tampón y las microesferas en un soporte magnético (cat. N^o: 12321D, Invitrogen, Carlsbad, USA) y dando vueltas al tubo (al menos 4 veces) de manera que las microesferas vayan de una pared a otra del tubo. Finalmente, se aspiró el volumen y se volvió a realizar un segundo lavado. Una vez lavadas, se resuspendieron en 30 μL de tampón de baja unión por muestra y se mantuvieron refrigeradas hasta su uso.
3. Posteriormente, se transfirieron 30 μL de las microesferas magnéticas de proteína A/G lavadas a cada tubo que contenía la mezcla del paso 1 en hielo y se incubaron a 4 °C en un

agitador rotatorio durante 90 minutos. De esta manera, el complejo ARN fragmentado unido al anticuerpo anti-m⁶A se une a las microesferas magnéticas, permitiendo la realización de los fututos lavados.

4. Terminado el tiempo de incubación, se colocó el tubo en el soporte magnético y se descartó el sobrenadante. En este paso el complejo ARN fragmentado unido al anticuerpo anti-m⁶A y a las microesferas magnéticas, permanece en la pared del tubo que se encuentra más próxima al soporte magnético.
5. Se realizaron dos lavados con 900 µL de tampón de baja unión. Cada lavado se rotó en una rueda giratoria durante 10 minutos a 4 °C.
6. Se realizaron dos lavados con 900 µL de tampón con alto contenido de sal. Cada lavado se mantuvo en un rotor giratorio durante 10 minutos a 4 °C.
7. Después de estos lavados extensivos, el complejo ARN fragmentado unido al anticuerpo anti-m⁶A se disoció de las microesferas magnéticas con la incorporación de 200 µL de tampón RLT suministrado por el RNeasy Mini Kit (cat. N°: 74106, Qiagen, Hilden, Alemania) y dejándolo actuar durante 2 minutos a temperatura ambiente. A continuación, se colocó el tubo en el soporte magnético y se recogió el sobrenadante, traspasándolo a un nuevo tubo en donde se incorporaron 400 µL de etanol al 100%. Se transfirió la mezcla a una columna de centrifugación RNeasy Mini Spin Column y se centrifugó a 12.000 rpm durante 1 minuto a 4 °C. Posteriormente, se realizaron dos lavados, uno con 500 µL del tampón RPE suministrado por el RNeasy Mini Kit y otro con 500 µL de etanol al 80%. Para eliminar el etanol residual, se realizó una nueva centrifugación a 13.000 rpm durante 5 minutos a 4 °C. Finalmente, se realizaron dos lavados con 7 µL de agua ultrapura libre de ARNasas suministrada por el kit. A la muestra de ARN inmunoprecipitado se le continuará llamado en adelante IP.

Tampones utilizados en la inmunoprecipitación

Tampón de baja unión (aproximadamente para 90 muestras):

- 20 mL de Tris-HCl a 1 M con pH 7,4 (cat. N°: 15568025, Invitrogen, Carlsbad, USA),
- 8 mL de NaCl a 5 M (cat. N°: 1064041000, Sigma-Aldrich, St. Louis, USA),
- 2 mL de NP-40 al 10% w/v (cat. N°: 85124, Thermo Scientific, Waltham, USA),
- 370 mL agua libre de ARNasas.

Tampón con alto contenido de sal (aproximadamente para 200 muestras):

- 20 mL de Tris-HCl a 1 M con pH 7,4,
- 80 mL de NaCl a 5 M,
- 40 mL de IGEPAL CA-630 al 10% w/v (cat. N°: I8896-50ML, Sigma-Aldrich, St. Louis, USA),
- 4 mL de EDTA al 100 mM (cat. N°: E5134, Sigma-Aldrich, St. Louis, USA),
- 20 mL de deoxicolato de sodio al 10% w/v (cat. N°: SRE0046-100G, Sigma-Aldrich, St. Louis, USA),
- 4 mL de SDS al 10% w/v (cat. N°: L4390, Sigma-Aldrich, St. Louis, USA),
- 232 mL agua libre de ARNasas.

3.5.3. Cuantificación de las muestras

Para poder determinar la concentración de ARN de la muestra IP, del control negativo y de la muestra INPUT, se realizó una cuantificación fluorimétrica con el Qubit® RNA HS Assay Kit tal y como se describe en el apartado 3.2.2.

3.5.4. Generación de la librería

La generación de la librería de las muestras IP e INPUT se llevó a cabo con el kit SMARTer Stranded Total RNA-Seq kit v2 – Pico Input Mammalian (cat. N°: 634412, Takara Clontech, Kyoto, Japan). Este kit presenta un flujo de trabajo adecuado para generar bibliotecas de ARN-seq compatibles con la plataforma Illumina. Resumidamente, el procedimiento utilizado para llevar a cabo la librería fue el siguiente: primero, el ARN se convirtió en ADN complementario (ADNc) (Protocolo A; este protocolo y los protocolos B, C, D, E y F indicados en este párrafo se pueden encontrar en el manual del kit) y luego se incorporaron los adaptadores para la secuenciación de Illumina (cada adaptador presenta códigos de barras específicos) mediante PCR utilizando solamente 5 ciclos (Protocolo B). Los productos de la PCR se purificaron (Protocolo C) y luego se eliminó el ADNc ribosómico (Protocolo D). Posteriormente, se realizó una amplificación con cebadores universales (Protocolo E) y, por último, los productos de la PCR se purificaron una vez más para producir la biblioteca de ADNc final (Protocolo F).

Para la generación de la biblioteca, se han seguido los protocolos propuestos por el fabricante, con las siguientes características particulares:

- En el Protocolo A, el ADNc se generó utilizando 50 ng de ARN, tanto para las muestras IP como para las muestras INPUT.
- En el Protocolo E, se escogieron diferentes números de ciclos para realizar la segunda PCR. En el caso de las muestras IP, se utilizaron 16 ciclos, mientras que para las muestras INPUT se utilizaron 12 ciclos.
- En el Protocolo F, las muestras fueron eluidas con 12 μL de Tris Buffer (provisto por el kit) en lugar de 20 μL porque el rendimiento esperado era bajo.

Cabe mencionar que, por cada muestra tumoral analizada, se realizaron dos librerías: una para la muestra IP y otra para la muestra INPUT, por lo que en total se generaron 40 librerías.

3.5.5. *Cuantificación de la librería y evaluación de la distribución de tamaños*

La cuantificación de las librerías se llevó a cabo utilizando el kit Qubit[®] 1X dsDNA HS Assay Kit, tal y como se describe en el apartado 3.3.2.2. Posteriormente, se realizó una dilución de cada librería a una concentración final de 1,50 ng/ μL , necesarios para el análisis en un Agilent 2100 Bioanalyzer (cat. N^o: G2939BA, Agilent, Santa Clara, USA).

La evaluación de la distribución del tamaño de la librería se realizó en un Agilent 2100 Bioanalyzer utilizando el kit (cat. N^o: 5067-4626, Agilent, Santa Clara, USA), siguiendo las instrucciones del fabricante. Los resultados del análisis se visualizaron con el programa 2100 Expert Software (cat. N^o: G2946CA, Agilent, Santa Clara, USA). La correcta generación de la librería debe producir en las muestras, una curva que abarque los 200-1.000 pb, con un máximo local en aproximadamente los 300-400 pb.

3.5.6. *Análisis de secuenciación masiva*

Para la secuenciación masiva, se realizaron diluciones de las librerías a 15 nM en 20 μL . Para ello, se utilizó la herramienta on-line: http://www.molbiol.edu.ru/eng/scripts/01_07.html, introduciendo los datos de la cuantificación obtenida mediante el Qubit[®] y el tamaño medio determinado por el Bioanalyzer.

La secuenciación masiva se llevó a cabo para las librerías generadas a partir de las muestras IP e INPUT en el secuenciador HiSeq 4.000 (Illumina) en la Unidad de Genómica del Centro Nacional de Investigaciones Cardiovasculares (CNIC, España). Después de la secuenciación, se

eliminaron los adaptadores y las lecturas de baja calidad. Las lecturas limpias se alinearon con el genoma de referencia versión GRCh38.91. Posteriormente, se identificaron los picos m⁶A, es decir, regiones que presentan una o varias metilaciones m⁶A. Los criterios establecidos para determinar que un pico fuese considerado como pico m⁶A fueron: por un lado, los picos debían encontrarse 5 veces más enriquecidos en la muestra IP que en la muestra INPUT y, por otro lado, las diferencias observadas debían ser estadísticamente significativas, presentando un valor de $p < 0,0001$.

Identificados los picos m⁶A, se procedió a la búsqueda de picos m⁶A consenso para cada subtipo molecular. En este caso, la condición fue que los picos m⁶A identificados como consenso debían estar presentes en todas las muestras analizadas para cada subtipo molecular. Posteriormente, se realizó un análisis de anotación de los picos m⁶A consenso mediante el programa informático HOMER (<http://homer.ucsd.edu/homer/>). Los picos m⁶A consenso de cada subtipo molecular se ordenaron en función del *peak score*, es decir, el valor de p transformado a $-\log_{10}(p)$ y se seleccionaron los primeros 20 picos m⁶A que presentaban el valor de *peak score* más elevado para continuar con los análisis.

3.5.7. Análisis de ontología génica y de rutas metabólicas

Para el top 20 de cada subtipo molecular, se llevó a cabo la clasificación funcional de la ontología génica (GO) utilizando la base de datos David Bioinformatics Resources 6.8 (<https://david.ncifcrf.gov/>) (191) y un análisis de la ruta metabólica de KEGG (*Kyoto Encyclopedia of genes and genomes*) incorporada en David. Los valores de p del test de fueron provistos por la base de datos David y fueron considerados como significativos cuando $p < 0,05$.

3.5.8. Análisis de expresión génica *in silico*

Con el fin de determinar si las posiciones m⁶A metiladas encontradas podrían tener un efecto en el cáncer de mama, se seleccionaron los cinco primeros genes de la lista top 20 de cada subtipo molecular para llevar a cabo un análisis de expresión génica *in silico*. Para ello, se utilizó la base de datos de la Universidad de California Santa Cruz (UCSC) (<https://xenabrowser.net/>) (192) que contiene los datos de expresión génica de muestras con cáncer de mama del *Cancer Genome Atlas* (TCGA). En total, contiene los datos de 119 muestras de tejido normal, 434 muestras con luminal A, 194 con luminal B, 67 con HER2+ y 142 con triple negativo. El valor de

p del estadístico t de Welch para las comparaciones entre la expresión del tejido normal y los distintos subtipos moleculares, como también, los diagramas de cajas fueron obtenidos a partir de dicha base de datos.

Para los genes que presentaron diferencias significativas en la expresión entre el tejido normal y el tejido tumoral de los distintos subtipos moleculares, se buscaron las posibles funciones de las proteínas codificadas por dichos genes mediante el uso de la base de datos del NCBI (<https://www.ncbi.nlm.nih.gov/>) (193) con el fin de evaluar la importancia que podrían tener en el proceso tumoral.

4. RESULTADOS

4.1. MUESTRAS ANALIZADAS EN EL ESTUDIO

En este estudio se analizaron muestras de tejido mamario tumoral y sano correspondientes a 35 pacientes con IBC-NST. Las características histopatológicas de las muestras fueron evaluadas y clasificadas en los distintos subtipos moleculares por anatompatólogos de amplia experiencia del Departamento de Anatomía Patológica del Hospital Universitario de Txagorritxu. Las muestras, así como los datos asociados a las mismas fueron proporcionadas por el Biobanco Vasco. La Tabla 4.1.1. contiene los datos asociados a las muestras.

Tabla 4.1.1. Datos asociados a las muestras de IBC-NST analizadas.

		Subtipos tumorales				
		Luminal A	Luminal B HER2-	Luminal B HER2+	HER2+	Triple negativo
Número de pacientes		7	7	8	6	7
Edad al diagnóstico (años)	Promedio DS.	52,57 6,35	66,29 0,91	59,88 8,69	65,83 8,38	55,14 7,93
Estado menopáusico al diagnóstico (%)	Si	-	85,71	87,50	100,00	71,43
	No	100,00	14,29	12,50	-	28,57
Recidivas (%)	Si	-	14,29	-	-	14,29
	No	100,00	85,71	100,00	100,00	85,71
TNM (% de muestras)	I	28,57	42,86	25,00	50,00	14,29
	IIA	28,57	28,57	50,00	-	42,85
	IIB	28,57	28,57	12,50	33,33	14,29
	IIIA	-	-	12,50	16,67	28,57
	IIIB	-	-	-	-	-
	IIIC	-	-	-	-	-
Grado histológico (% de muestra)	IV	14,29	-	-	-	-
	I	28,57	-	-	-	-
	II	71,43	42,86	62,50	16,67	14,29
Vascular linfática (%)	III	-	57,14	37,50	83,33	85,71
	Si	14,29	14,29	25,00	-	42,86
Multicentricidad del tumor (%)	No	85,71	85,71	75,00	100,00	57,14
	Si	14,29	14,29	12,50	16,67	-
Necrosis tumoral (%)	No	85,71	85,71	87,50	83,33	100
	Si	28,57	14,29	37,50	66,67	85,71
Receptores estrógenos (%)	No	71,43	85,71	62,50	33,33	14,29
	Negativo	-	-	12,50	100,00	100,00
	Leve	-	-	-	-	-
	Moderado	-	-	75,00	-	-
Receptores progesterona (%)	Intenso	100,00	100,00	12,50	-	-
	Negativo	-	-	12,50	100,00	100,00
	Leve	-	-	12,50	-	-
	Moderado	28,57	42,86	25,00	-	-
HER2 (FISH) (%)	Intenso	71,43	57,14	50,00	-	-
	Sin dato	-	-	-	-	28,57
	<2	100,00	100,00	-	-	71,43
	>2	-	-	100,00	100,00	-

		Subtipos tumorales				
		Luminal A	Luminal B HER2-	Luminal B HER2+	HER2+	Triple negativo
HER2 (IHQ) (%)	0	42,86	42,86	-	-	57,14
	+1	28,57	28,57	-	-	28,57
	+2	28,57	28,57	50,00	33,33	14,29
	+3	-	-	50,00	66,67	-
Índice Ki-67	<10	14,28	-	-	-	-
	10-25	85,71	28,57	25,00	-	14,29
	>25	-	71,43	75,00	100,00	85,71
CK 19 (%)	Positiva	100,00	100,00	100,00	83,33	100,00
	Negativa	-	-	-	-	-
	Sin dato	-	-	-	16,67	-
CK 5/6 (%)	Positiva	-	-	12,50	33,33	85,71
	Negativa	100,00	100,00	87,50	50,00	14,29
	Sin dato	-	-	-	16,67	-

Continuación de la Tabla 4.1.1.

Todas las muestras analizadas se caracterizaban por ser tumores primarios. La mayor parte de las muestras no presentaban multicentricidad y únicamente dos de ellas eran recidivas y se correspondían a los subtipos luminal B HER2- y triple negativo. Los estadios TNM I y II fueron en general los más comunes. En cuanto al grado histológico, la mayor parte de las muestras eran del grado I y II, aunque las muestras provenientes de los subtipos HER2+ y triple negativo mostraron un grado histológico mayor que el resto de los subtipos. El subtipo triple negativo presentó un elevado porcentaje de muestras en donde el tumor ya había invadido el sistema linfático y vascular. Además, los subtipos HER2+ y triple negativo mostraron un elevado porcentaje de muestras con necrosis tumoral.

En cuanto a los marcadores ER, PR y HER2, en las muestras estaban expresados aquellos que eran característicos al subtipo molecular al que pertenecían. Por otra parte, el marcador de proliferación celular Ki-67 presentó una expresión intermedia-alta en la mayoría de los subtipos, a excepción del subtipo luminal A. Todas las muestras fueron positivas para el marcador CK 19 y el subtipo triple negativo presentó una expresión elevada de CK 5/6 que permite clasificar este subtipo en subgrupos.

4.2. ANÁLISIS DE LA METILACIÓN m⁶A

4.2.1. Determinación del porcentaje de metilación m⁶A global

Para explorar el papel potencial de la modificación epigenética m⁶A en el cáncer de mama, primeramente, se llevó a cabo la determinación del porcentaje de metilación m⁶A en el ARN

total de los tejidos tumorales y de sus correspondientes tejidos sanos mediante el kit EpiQuik m⁶A RNA Methylation Quantification Kit (Colorimetric). Este análisis se llevó a cabo únicamente en muestras de 20 pacientes debido a que es el número máximo de muestras que se puede analizar simultáneamente en la misma placa en la que se obtiene la curva estándar. Usando esta estrategia de cuantificación colorimétrica de m⁶A se observó que el porcentaje de metilación se encontraba disminuido en los tejidos tumorales en comparación con los tejidos sanos ($0,367 \pm 0,094$ en las muestras tumorales y $0,473 \pm 0,075$ en las muestras sanas) (Figura 4.2.1A) y que dicha diferencia era estadísticamente significativa cuando se realizó la prueba de U de Mann-Whitney ($p = 0,001$).

Posteriormente, el porcentaje de metilación m⁶A global se evaluó teniendo en cuenta los distintos subtipos moleculares (luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo). Los porcentajes de metilación m⁶A variaron de 0,429 - 0,536% para el tejido mamario sano y de 0,279 - 0,420% para el tejido tumoral (Tabla 4.2.1 y Figura 4.2.1B-F). En todos los casos, el porcentaje de metilación fue menor en el tejido tumoral que en el tejido sano. El subtipo luminal A presentó el porcentaje de metilación más elevado en las muestras sanas y el porcentaje de metilación más bajo en las muestras tumorales (Tabla 4.2.1). Este fue el único subtipo que presentó diferencias significativas cuando se compararon los niveles de metilación entre los tejidos sanos y tumorales ($p = 0,043$) (Tabla 4.2.1). También se comparó el porcentaje de metilación obtenido en las muestras tumorales de los diferentes subtipos mediante el test Kruskal-Wallis y no se encontraron diferencias significativas ($p = 0,087$) (Tabla 4.2.1 y Figura 4.2.1G).

Tabla 4.2.1. Promedio del porcentaje de metilación m⁶A global con su desviación estándar (DS) en el tejido tumoral y en su correspondiente tejido sano, así como los valores de p obtenidos para las pruebas U de Mann-Whitney y H de Kruskal-Wallis. Los valores de $p < 0,05$ son considerados como significativos (indicados en negrita).

Pruebas estadísticas		Luminal A	Luminal B HER2-	Luminal B HER2+	HER2+	Triple negativo
Promedio \pm DS	Sano	0,536 \pm 0,045	0,444 \pm 0,055	0,439 \pm 0,099	0,429 \pm 0,078	0,518 \pm 0,036
	Tumoral	0,279 \pm 0,164	0,420 \pm 0,039	0,398 \pm 0,038	0,333 \pm 0,028	0,405 \pm 0,081
U de Mann-Whitney (p)		0,043	0,564	0,564	0,149	0,083
H de Kruskal-Wallis (p)		0,087				

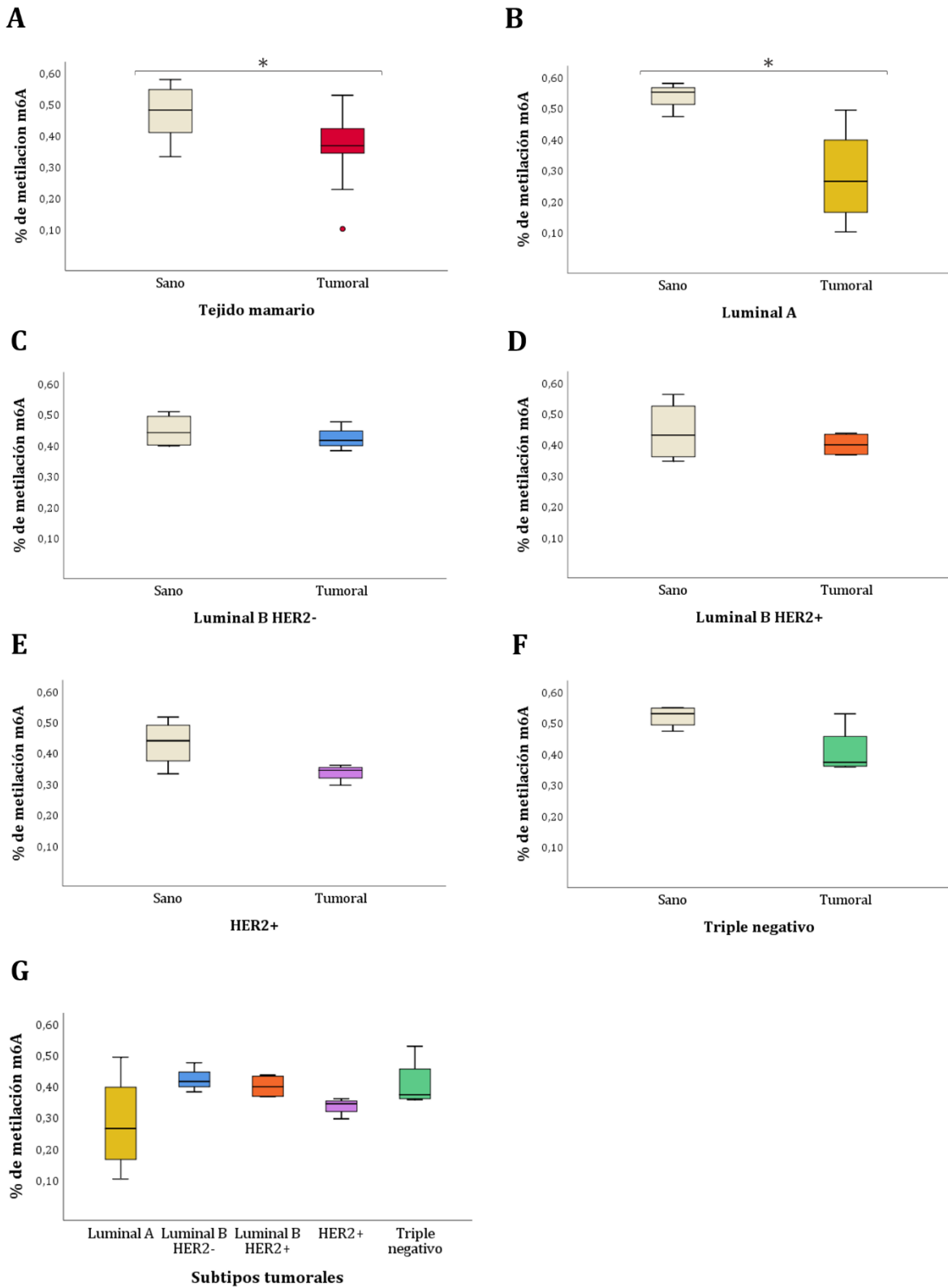


Figura 4.2.1. Porcentaje de metilación m⁶A global entre las muestras tumorales y sus correspondientes tejidos sanos: **A-** Todas las muestras analizadas, **B-** Subtipo luminal A, **C-** Subtipo luminal B HER2-, **D-** Subtipo luminal B HER2+, **E-** Subtipo HER2+ y **F-** Subtipo triple negativo. **G-** Porcentaje de metilación m⁶A global entre los diferentes subtipos tumorales. * $P < 0,05$. Los círculos son valores atípicos, es decir, valores que se encuentran fuera de los siguientes rangos: 3^{er} cuartil+1,5*rango intercuartílico y 1^{er} cuartil+1,5*rango intercuartílico.

4.2.2. Análisis de expresión génica

La metilación m⁶A es un proceso dinámico y reversible que se encuentra catalizado por las metiltransferasas y desmetilasas. Por tal motivo, se procedió a evaluar la expresión génica mediante una RT-qPCR de los principales componentes involucrados en el proceso de metilación (*METTL3*, *METTL14* y *WTAP*) y desmetilación (*FTO* y *ALKBH5*), con el fin de determinar si la disminución del porcentaje de metilación m⁶A global en las muestras tumorales era causado por la desregulación de dichos componentes. Este análisis se llevó a cabo en las muestras de las 35 pacientes con cáncer de mama disponibles para el estudio.

Previo al análisis de expresión, se evaluó la eficiencia de las parejas de cebadores. El análisis de eficiencia es de suma importancia ya que la expresión de los genes diana (*METTL3*, *METTL14*, *WTAP*, *FTO* y *ALKBH5*) será normalizada por los genes *housekeeping* (*GUSB*, *GAPDH* y *BACT*) utilizando los valores del Cq. Por lo tanto, los genes dianas y los *housekeeping* deben amplificarse con eficiencias similares para que la normalización sea precisa. En todas las parejas de cebadores se obtuvieron eficiencias >90% (Figura 4.2.2-4.2.4 y Tabla 4.2.2). Además,

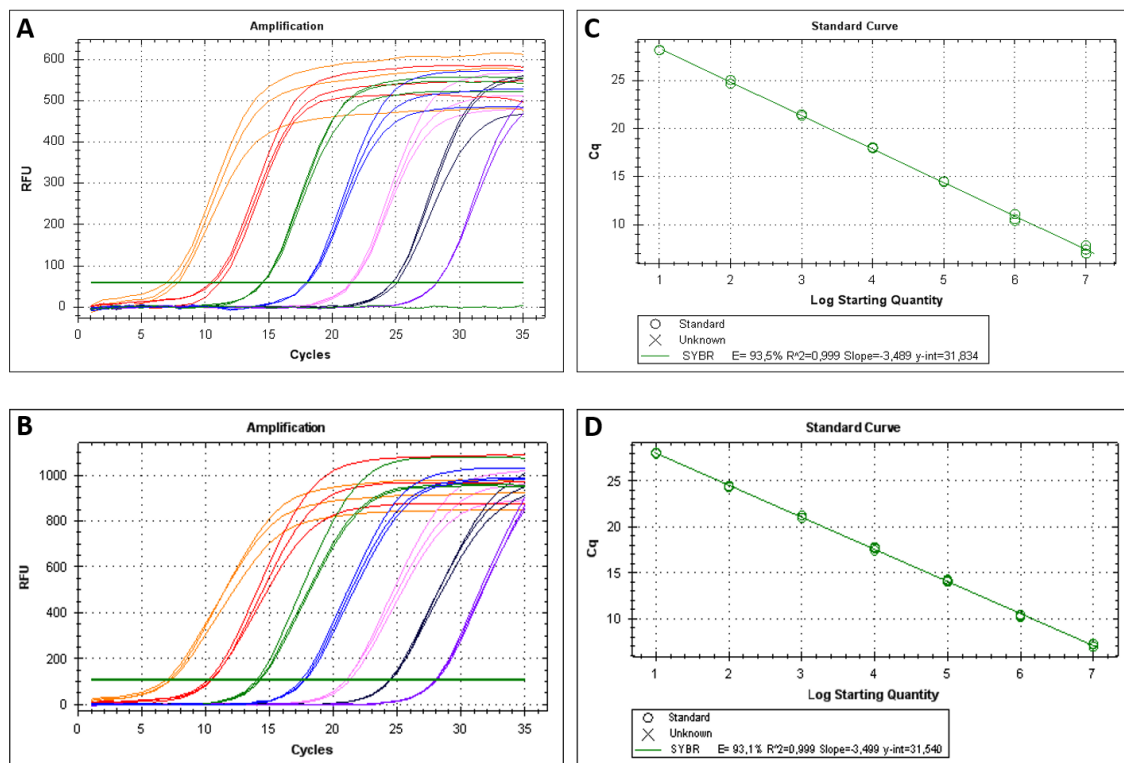


Figura 4.2.2. Análisis de eficiencia de las parejas de cebadores utilizados en la expresión de los genes que codifican las desmetilasas *FTO* y *ALKBH5*. Curvas de amplificación de las diluciones realizadas a diferentes números de copias: **A-** *FTO*, **B-** *ALKBH5*. Recta obtenida a partir de los valores Cq de las curvas de amplificación: **C-** *FTO*, **D-** *ALKBH5*. E= eficiencia; R² = coeficiente de correlación; Slope = pendiente; Y-int = valor de intersección con el eje y. Concentraciones: Naranja 1x10⁷ copias/μL; Rojo: 1x10⁶ copias/μL; Verde: 1x10⁵ copias/μL; Azul: 1x10⁴ copias/μL; Rosa: 1x10³ copias/μL; Negro: 1x10² copias/μL; Lila: 1x10¹ copias/μL.

las diferencias en la eficiencia entre los genes diana y los *housekeeping* fueron <10%, lo que indica la idoneidad de los pares de cebadores del presente estudio para realizar el análisis de expresión génica.

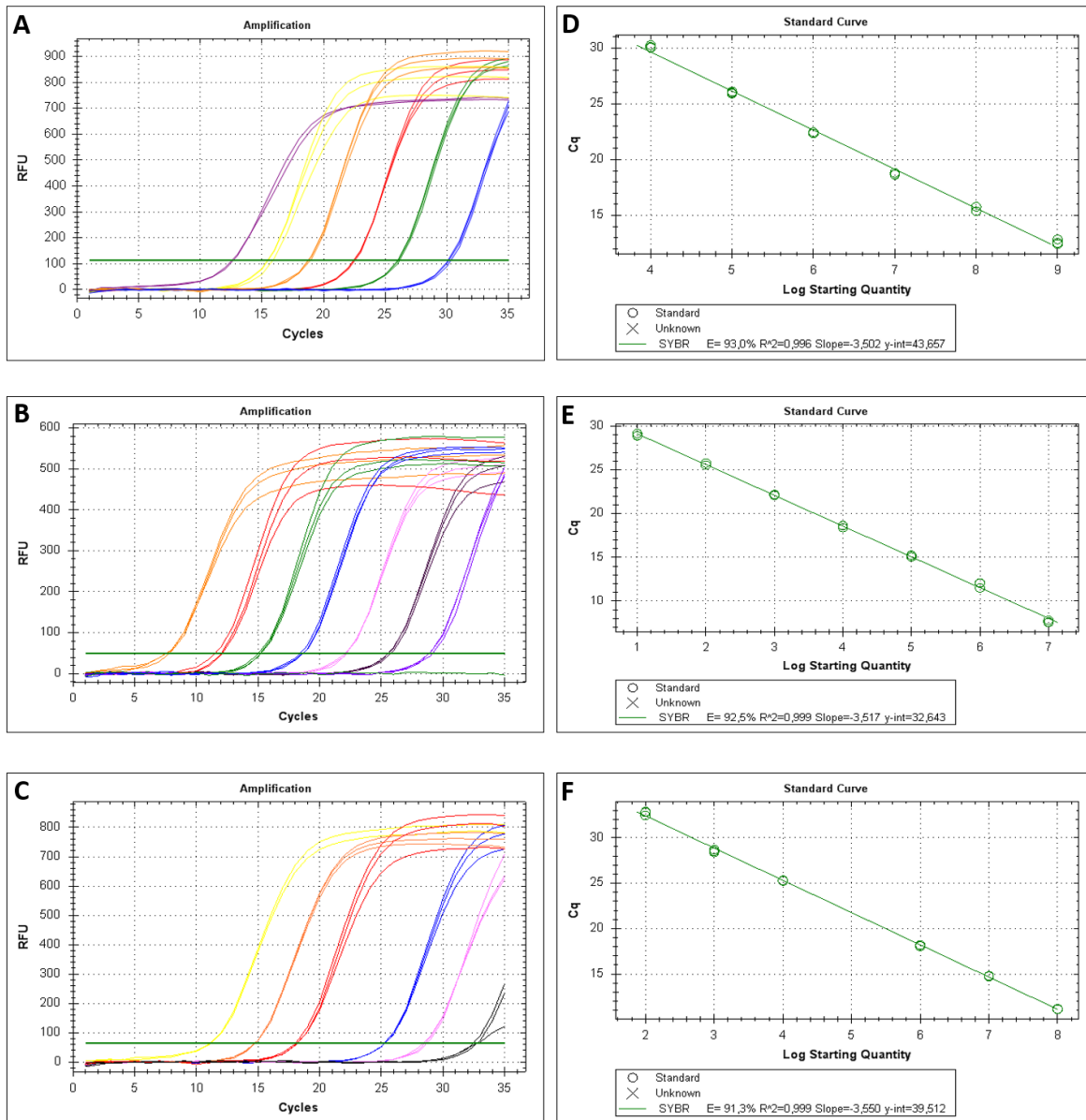


Figura 4.2.3. Análisis de eficiencia de las parejas de cebadores utilizados en la expresión de los genes que codifican las metiltransferasas *METTL3*, *METTL14* y *WTAP*. Curvas de amplificación de las diluciones realizadas a distintos números de copias: **A-** *METTL3*, **B-** *METTL14*, **C-** *WTAP*. Recta obtenida a partir de los valores Cq de las curvas de amplificación: **D-** *METTL3*, **E-** *METTL14*, **F-** *WTAP*. E = eficiencia; R² = coeficiente de correlación; Slope = pendiente; Y-int = valor de intersección con el eje y. Concentraciones: Violeta: 1 x10⁹ copias/μL; Amarillo: 1x10⁸ copias/μL; Naranja 1x10⁷ copias/μL; Rojo: 1x10⁶ copias/μL; Verde: 1x10⁵ copias/μL; Azul: 1x10⁴ copias/μL; Rosa: 1x10³ copias/μL; Negro: 1x10² copias/μL; Lila: 1x10¹ copias/μL.

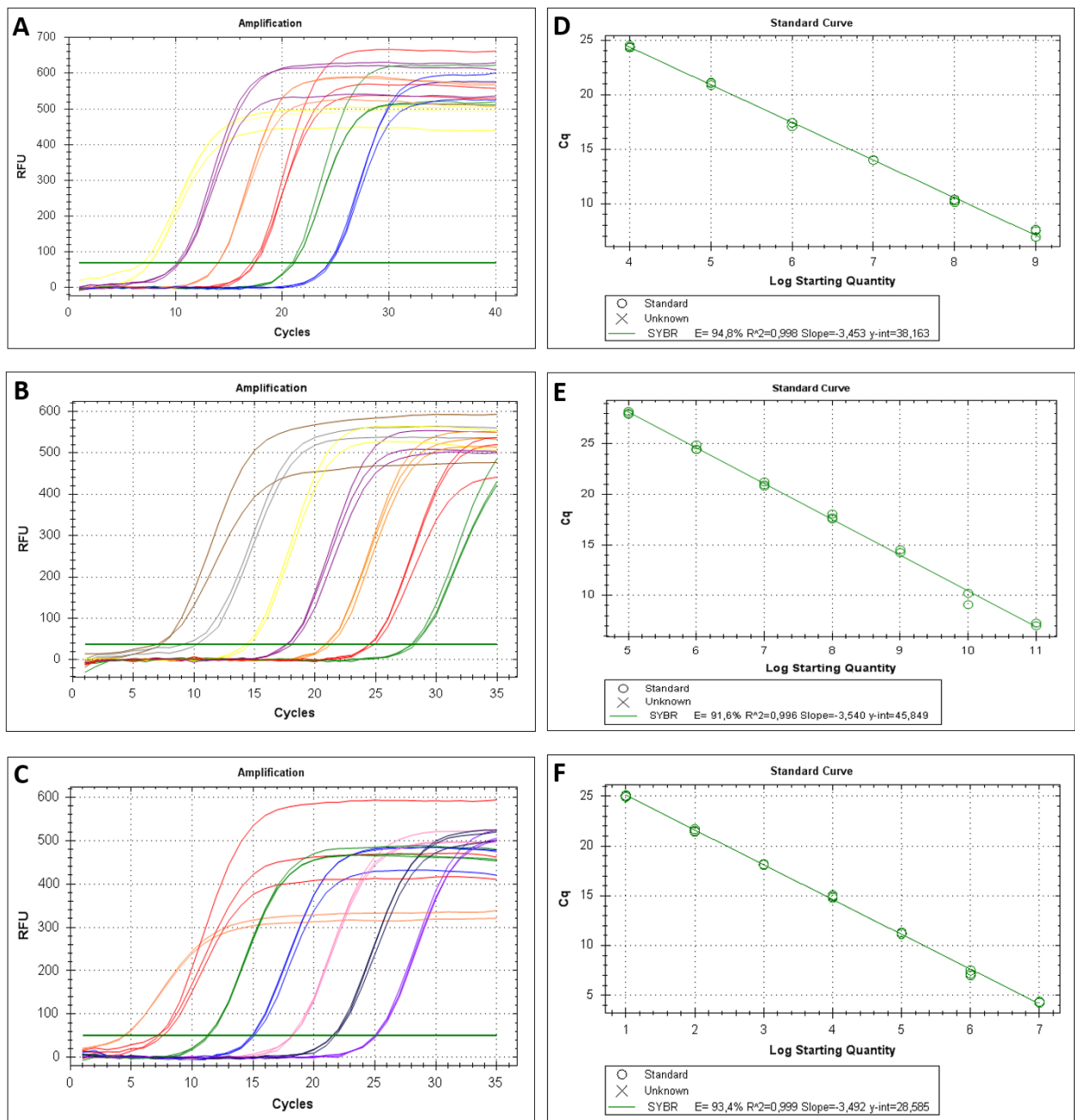


Figura 4.2.4. Análisis de eficiencia de las parejas de cebadores utilizados en la expresión de los genes *housekeeping* *GUSB*, *GAPDH* y *ACTB*. Curvas de amplificación de las diluciones realizadas a distintos números de copias: **A-** *GUSB*, **B-** *GAPDH*, **C-** *ACTB*. Recta obtenida a partir de los valores Cq de las curvas de amplificación: **D-** *GUSB*, **E-** *GAPDH*, **F-** *ACTB*. E = eficiencia; R^2 = coeficiente de determinación; *Slope* = pendiente; Y-int = valor de intersección con el eje y. Concentraciones: Marrón: 1×10^{11} copias/ μL ; Gris: 1×10^{10} copias/ μL ; Amarillo: 1×10^9 copias/ μL ; Violeta: 1×10^8 copias/ μL ; Naranja 1×10^7 copias/ μL ; Rojo: 1×10^6 copias/ μL ; Verde: 1×10^5 copias/ μL ; Azul: 1×10^4 copias/ μL ; Rosa: 1×10^3 copias/ μL ; Negro: 1×10^2 copias/ μL ; Lila: 1×10^1 copias/ μL .

Tabla 4.2.2. Resumen de los valores de eficiencia, coeficiente de correlación (R^2), pendiente y la intersección con el eje Y, obtenidos para los genes *METTL3*, *METTL14*, *WTAP*, *FTO*, *ALKBH5*, *GUSB*, *GADPH* y *ACTB*.

Gen	Eficiencia (%)	Coefficiente de correlación (R^2)	Pendiente	Intersección con el eje Y
<i>METTL3</i>	93,00	0,996	-3,50	43,66
<i>METTL14</i>	92,50	0,999	-3,52	32,64
<i>WTAP</i>	91,30	0,999	-3,55	39,51
<i>FTO</i>	93,50	0,999	-3,49	31,83
<i>ALKBH5</i>	93,10	0,999	-3,50	31,54
<i>GUSB</i>	94,80	0,998	-3,45	38,16
<i>GADPH</i>	91,60	0,996	-3,54	45,84
<i>ACTB</i>	93,40	0,999	-3,49	28,59

En cuanto al análisis de expresión génica, encontramos que la expresión de las metiltransferasas y desmetilasas analizadas estaba desregulada en las muestras tumorales y, en todos los casos, su expresión fue significativamente menor que en las muestras de tejidos sanos (Figura 4.2.5 y Tabla 4.2.3). Posteriormente, se analizó la expresión génica entre las muestras tumorales y los tejidos sanos teniendo en cuenta los diferentes subtipos moleculares. En la mayoría de los subtipos, la expresión fue más baja en las muestras de tejido tumoral que en los tejidos sanos (Tabla 4.2.4). Además, los niveles de expresión de los genes analizados

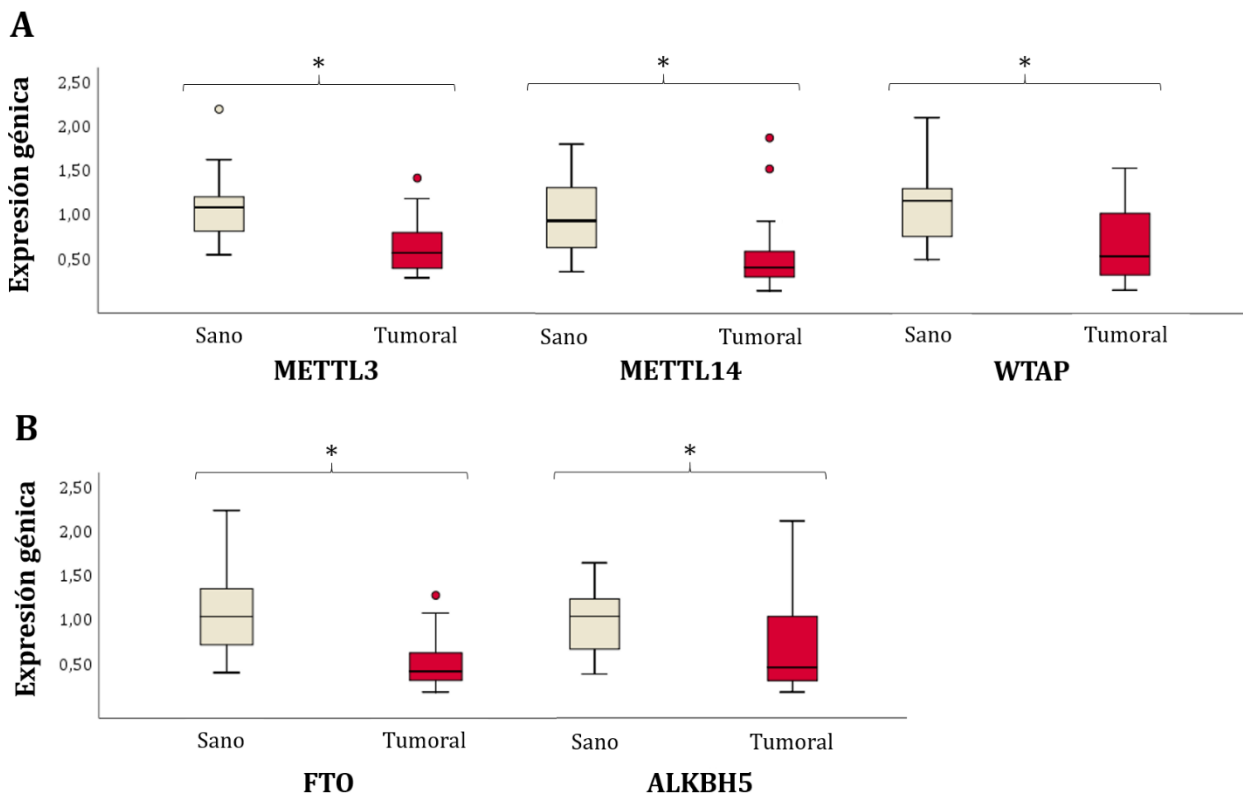


Figura 4.2.5. Análisis de expresión génica entre las muestras de tejidos tumorales y de sus correspondientes tejidos sanos. **A-** Metiltransferasas (*METTL3*, *METTL14* y *WTAP*); **B-** Desmetilasas (*FTO* y *ALKBH5*). * $P < 0,05$. Los círculos son valores atípicos.

mostraron diferencias significativas según el subtipo molecular. En los genes de las metiltransferasas, *METTL3* presentó diferencias significativas en los subtipos luminal A ($p = 0,006$), luminal B HER2- ($p = 0,018$) y triple negativo ($p = 0,004$) (Figura 4.2.6 y Tabla 4.2.4). *METTL14* presentó diferencias en todos los subtipos analizados con excepción del subtipo luminal B HER2+ (Figura 4.2.6 y Tabla 4.2.4) y *WTAP* tuvo diferencias en los subtipos luminal A ($p = 0,002$) y HER2+ ($p = 0,004$) (Figura 4.2.6 y Tabla 4.2.4). Por otro parte, en los genes desmetilasas, *FTO* presentó diferencias significativas en todos los subtipos (Figura 4.2.7 y Tabla 4.2.4), mientras que las diferencias en la expresión de *ALKBH5* fueron en los subtipos luminal A ($p = 0,002$) y HER2+ ($p = 0,004$) (Figura 4.2.7 y Tabla 4.2.4). La Figura 4.2.8 presenta un resumen de los resultados obtenidos en el análisis de expresión génica.

Tabla 4.2.3. Promedio de la expresión génica relativa para los genes *METTL3*, *METTL14*, *WTAP*, *FTO* y *ALKBH5* con su desviación estándar (DS) en el tejido tumoral y su correspondiente tejido mamario sano, y el valor de p obtenido para la prueba U de Mann-Whitney. Los valores de $p < 0,05$ son considerados como significativos (indicados en negrita).

Pruebas estadísticas		<i>METTL3</i>	<i>METTL14</i>	<i>WTAP</i>	<i>FTO</i>	<i>ALKBH5</i>
Promedio \pm DS	Sano	1,047 \pm 0,330	0,980 \pm 0,340	1,054 \pm 0,346	1,104 \pm 0,493	0,979 \pm 0,382
	Tumoral	0,593 \pm 0,267	0,474 \pm 0,350	0,646 \pm 0,426	0,455 \pm 0,263	0,660 \pm 0,523
U de Mann-Whitney (p)		<0,001	<0,001	<0,001	<0,001	0,001

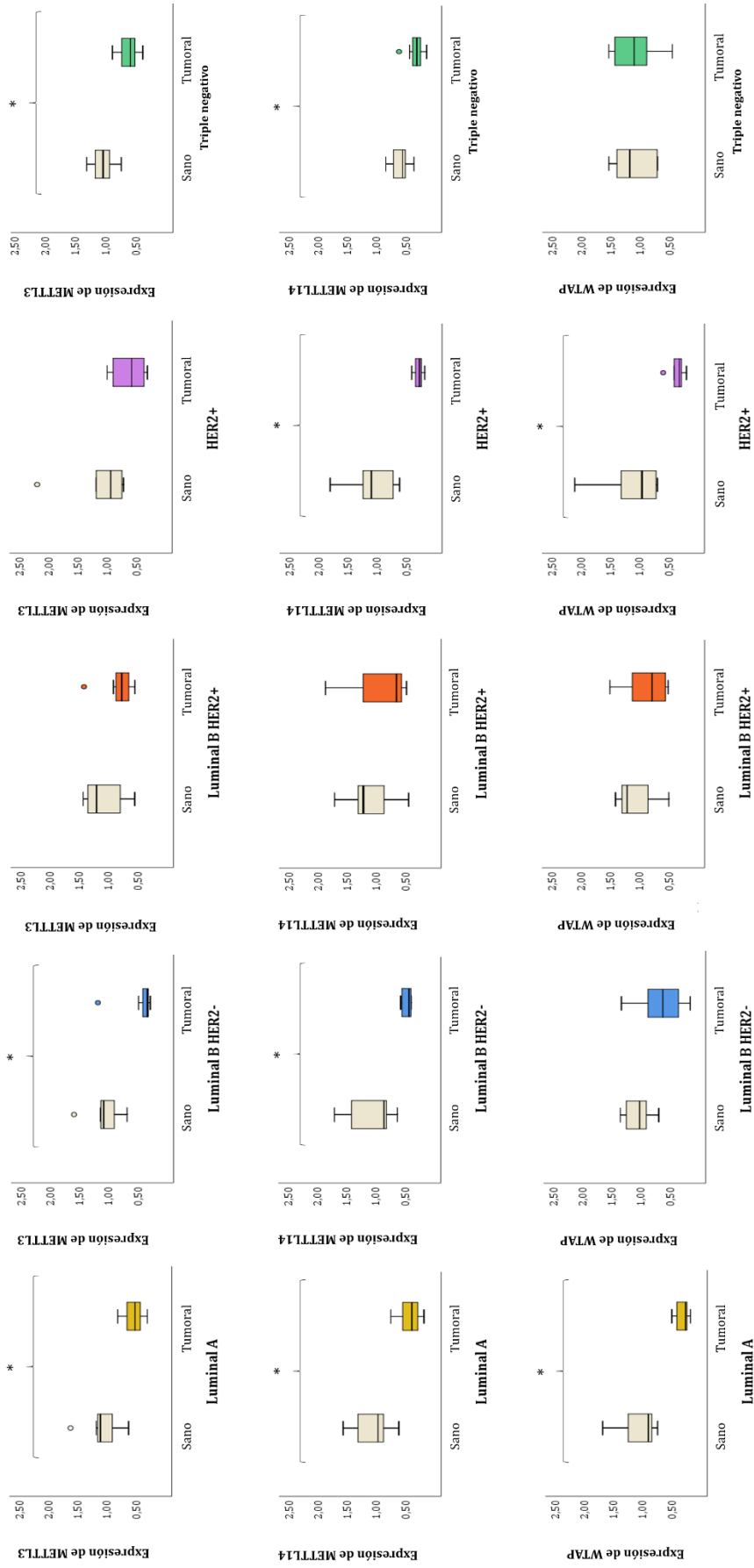


Figura 4.2.6. Expresión génica relativa de las metiltransferasas *METTL3*, *METTL14* y *WTAP* para las muestras tumorales y sus correspondientes tejidos sanos en los diferentes subtipos moleculares. * $P < 0,05$. Los círculos son valores atípicos.

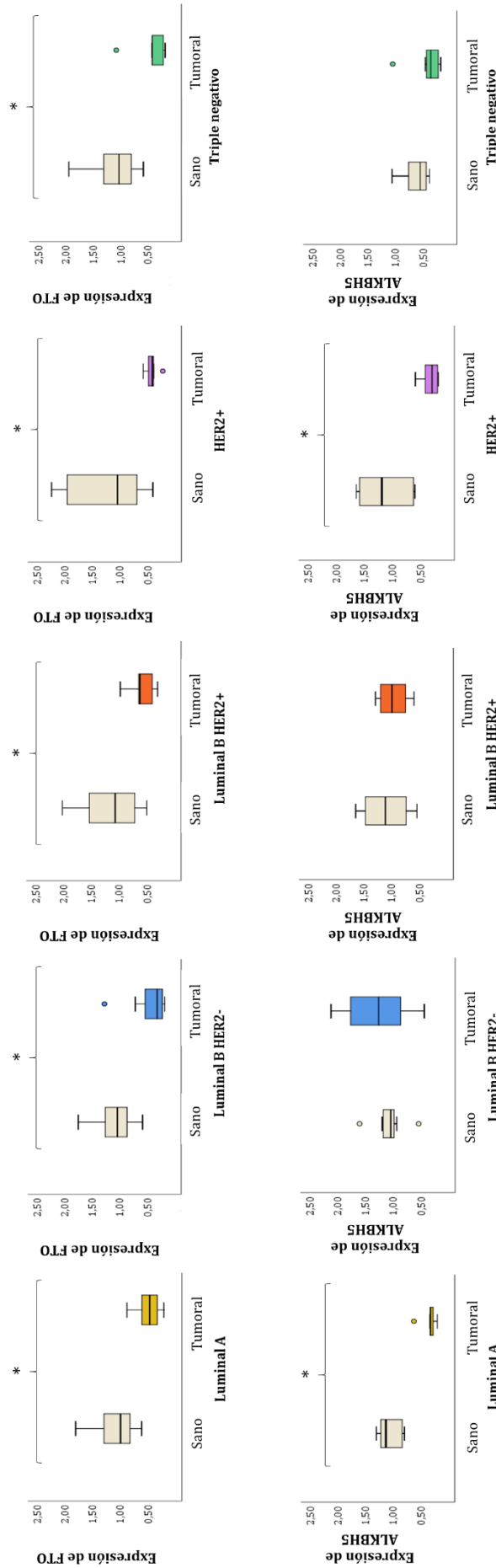


Figura 4.2.7. Expresión génica relativa de las desmetilasas *FTO* y *ALKBH5* para las muestras tumorales y sus correspondientes tejidos sanos en los diferentes subtipos moleculares. * $P < 0,05$. Los círculos son valores atípicos.

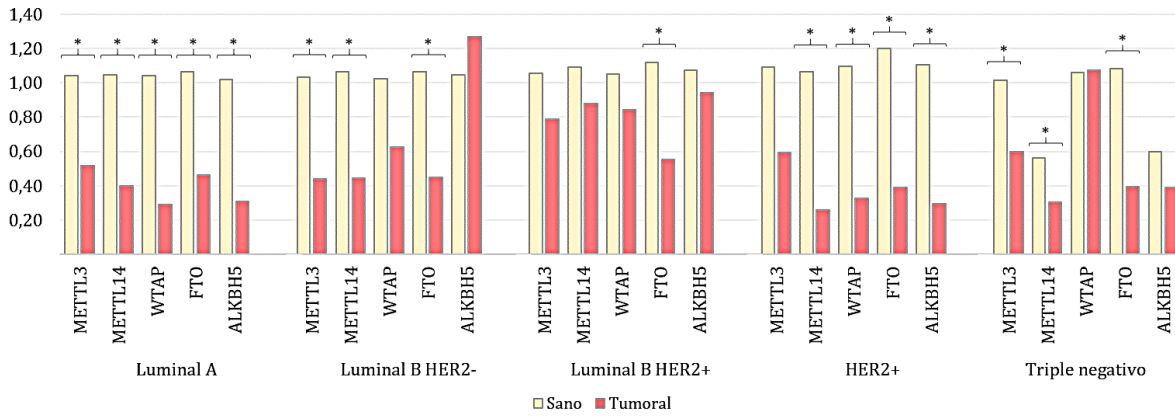


Figura 4.2.8. Resumen de los niveles de expresión de los genes *METTL3*, *METTL14*, *WTAP*, *FTO* y *ALKBH5* obtenidos en los tejidos tumorales y sus correspondientes tejidos sanos para los diferentes subtipos moleculares. * $P < 0,05$.

Tabla 4.2.4. Promedio de la expresión génica relativa para los genes *METTL3*, *METTL14*, *WTAP*, *FTO* y *ALKBH5* con su desviación estándar (DS) en el tejido tumoral y su correspondiente tejido sano para cada subtipo, y los valores de p obtenidos para las pruebas U de Mann-Whitney y H de Kruskal-Wallis. Los valores de $p < 0,05$ son considerados como significativos (indicados en negrita).

Gen	Pruebas estadísticas		Luminal A	Luminal B HER2-	Luminal B HER2+	HER2+	Triple negativo
		Promedio ± DS	Sano Tumoral				
<i>METTL3</i>	Promedio ± DS	Sano Tumoral	1,042 ± 0,316 0,515 ± 0,175	1,033 ± 0,287 0,439 ± 0,322	1,056 ± 0,335 0,790 ± 0,274	1,091 ± 0,560 0,595 ± 0,277	1,017 ± 0,195 0,599 ± 0,181
	U de Mann-Whitney (p)		0,006	0,018	0,141	0,055	0,004
	H de Kruskal-Wallis (p)		0,056				
<i>METTL14</i>	Promedio ± DS	Sano Tumoral	1,048 ± 0,342 0,399 ± 0,209	1,066 ± 0,412 0,443 ± 0,085	1,092 ± 0,419 0,877 ± 0,515	1,067 ± 0,422 0,260 ± 0,079	0,560 ± 0,168 0,304 ± 0,156
	U de Mann-Whitney (p)		0,004	0,002	0,462	0,004	0,022
	H de Kruskal-Wallis (p)		0,002				
<i>WTAP</i>	Promedio ± DS	Sano Tumoral	1,042 ± 0,337 0,290 ± 0,119	1,025 ± 0,240 0,624 ± 0,415	1,052 ± 0,315 0,841 ± 0,357	1,096 ± 0,547 0,324 ± 0,137	1,059 ± 0,373 1,074 ± 0,401
	U de Mann-Whitney (p)		0,002	0,064	0,208	0,004	0,749
	H de Kruskal-Wallis (p)		0,001				
<i>FTO</i>	Promedio ± DS	Sano Tumoral	1,064 ± 0,406 0,464 ± 0,232	1,063 ± 0,393 0,447 ± 0,397	1,119 ± 0,553 0,554 ± 0,211	1,202 ± 0,744 0,391 ± 0,117	1,082 ± 0,476 0,394 ± 0,308
	U de Mann-Whitney (p)		0,006	0,025	0,036	0,025	0,009
	H de Kruskal-Wallis (p)		0,420				
<i>ALKBH5</i>	Promedio ± DS	Sano Tumoral	1,021 ± 0,219 0,309 ± 0,135	1,046 ± 0,318 1,270 ± 0,665	1,075 ± 0,416 0,942 ± 0,255	1,105 ± 0,508 0,296 ± 0,158	0,598 ± 0,253 0,388 ± 0,299
	U de Mann-Whitney (p)		0,002	0,338	0,462	0,004	0,063
	H de Kruskal-Wallis (p)		<0,001				

También se evaluó si existían diferencias significativas entre muestras tumorales de los distintos subtipos mediante la prueba estadística H de Kruskal-Wallis. La expresión de *METTL3* y *FTO* fue similar en todos los subtipos y no mostró diferencias significativas (Figura 4.2.9, Tabla 4.2.4). Por el contrario, la expresión de los genes *METTL14*, *WTAP* y *ALKBH5* mostró diferencias entre algunos subtipos como se muestra en la Tabla 4.2.5. La expresión de *METTL14* en el subtipo luminal B HER2+ presentó diferencias con respecto al resto de subtipos ya que presentaba una mayor expresión (Figura 4.2.9, Tabla 4.2.4 y 4.2.5). También, presentó diferencias entre los subtipos luminal B HER2- y HER2+. La expresión de *WTAP*, presentó diferencias, por un lado, en el subtipo triple negativo con respecto a los subtipos luminal A, luminal B HER2- y HER2+ y, por otro lado, el subtipo luminal B HER2+ con respecto a los subtipos luminal A y HER2+ (Figura 4.2.9, Tabla 4.2.4 y 4.2.5). Estos dos subtipos (triple negativo y luminal B HER2+) tienen una mayor expresión de *WTAP* que el resto de subtipos. Finalmente, la expresión de *ALKBH5* fue más alta en los subtipos luminal B HER2- y luminal B HER2+ que en los otros subtipos tumorales (Figura 4.9.9 y Tabla 4.2.4 y 4.2.5).

Tabla 4.2.5. Valores de *p* obtenidos tras realizar las pruebas estadísticas U de Mann-Whitney con el fin de determinar qué subtipos tumorales presentaban las diferencias estadísticas observadas en la prueba H de Kruskal-Wallis. Los valores de *p* <0,05 son considerados como significativos (indicados en negrita).

Gen	Subtipos	Luminal A	Luminal B HER2-	HER2+	Triple negativo
<i>METTL14</i>	Luminal B HER2+	0,028	0,011	0,002	0,004
	Luminal A	-	0,338	0,253	0,565
	Luminal B HER2-	-	-	0,010	0,085
	HER2+	-	-	-	0,668
<i>WTAP</i>	Luminal B HER2+	0,001	0,298	0,007	0,298
	Luminal A	-	0,110	0,568	0,003
	Luminal B HER2-	-	-	0,153	0,048
	HER2+	-	-	-	0,004
<i>ALKBH5</i>	Luminal B HER2+	0,002	0,247	0,002	0,008
	Luminal A	-	0,004	0,775	0,749
	Luminal B HER2-	-	-	0,007	0,006
	TNBC	-	-	-	0,568

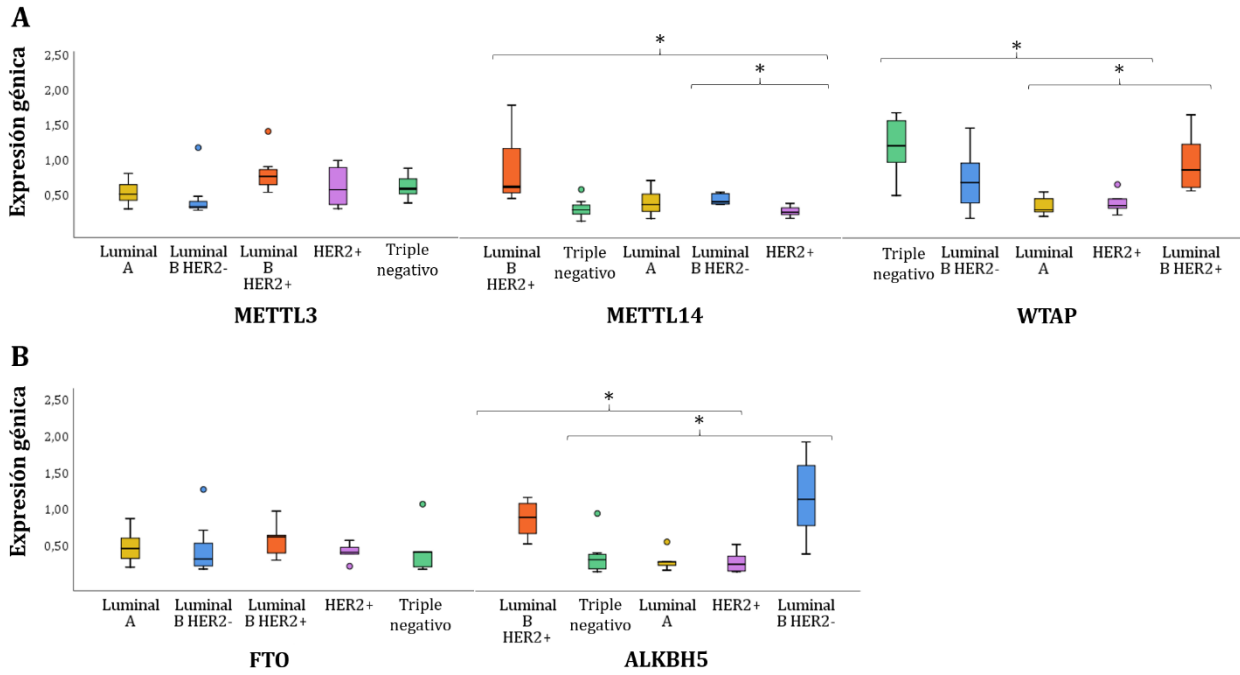


Figura 4.2.9. Niveles de expresión génica entre las muestras tumorales de los diferentes subtipos moleculares. **A-** Metiltransferasas (*METTL3*, *METTL14* y *WTAP*); **B-** Desmetilasas (*ALKBH5* y *FTO*). **P* < 0,05. Los círculos son valores atípicos.

4.2.3. Análisis de correlación

Dado que las metiltransferasas y desmetilasas están involucradas en la modificación m⁶A, se esperaría que estuvieran correlacionadas entre sí. Por tal motivo, evaluamos cómo se correlacionan a nivel de expresión génica utilizando la prueba estadística de correlación de Pearson. Como se muestra en la Figura 4.2.10, se observa una correlación significativa entre la expresión génica de *FTO* y *METTL14* ($r = 0,40$, $p = 0,017$), *FTO* y *WTAP* ($r = 0,48$, $p = 0,003$), *METTL14* y *ALKBH5* ($r = 0,45$, $p = 0,006$) y *METTL14* y *METTL3* ($r = 0,53$, $p = 0,001$) en los tejidos tumorales de las 35 pacientes con cáncer de mama sin distinguir entre subtipos. Posteriormente, se examinaron las correlaciones teniendo en cuenta los distintos subtipos moleculares. En este caso, se encontraron fuertes correlaciones entre metiltransferasas y desmetilasas solo en algunos subtipos moleculares (Figura 4.2.10). Además, todos los subtipos excepto HER2+, se caracterizaron por la correlación de la expresión génica alterada de *FTO* con *METTL14* y *ALKBH5* en los subtipos luminal A y triple negativo, y con *WTAP* en los subtipos luminal B HER2- y luminal B HER2+. En el caso del subtipo triple negativo, se encontró una correlación adicional entre la expresión génica de *METTL14* y *ALKBH5*, mientras que en los subtipos luminal B HER2- y luminal B HER2+, se encontraron correlaciones adicionales entre *WTAP* y *METTL14*, y *METTL3* y *METTL14*, respectivamente.

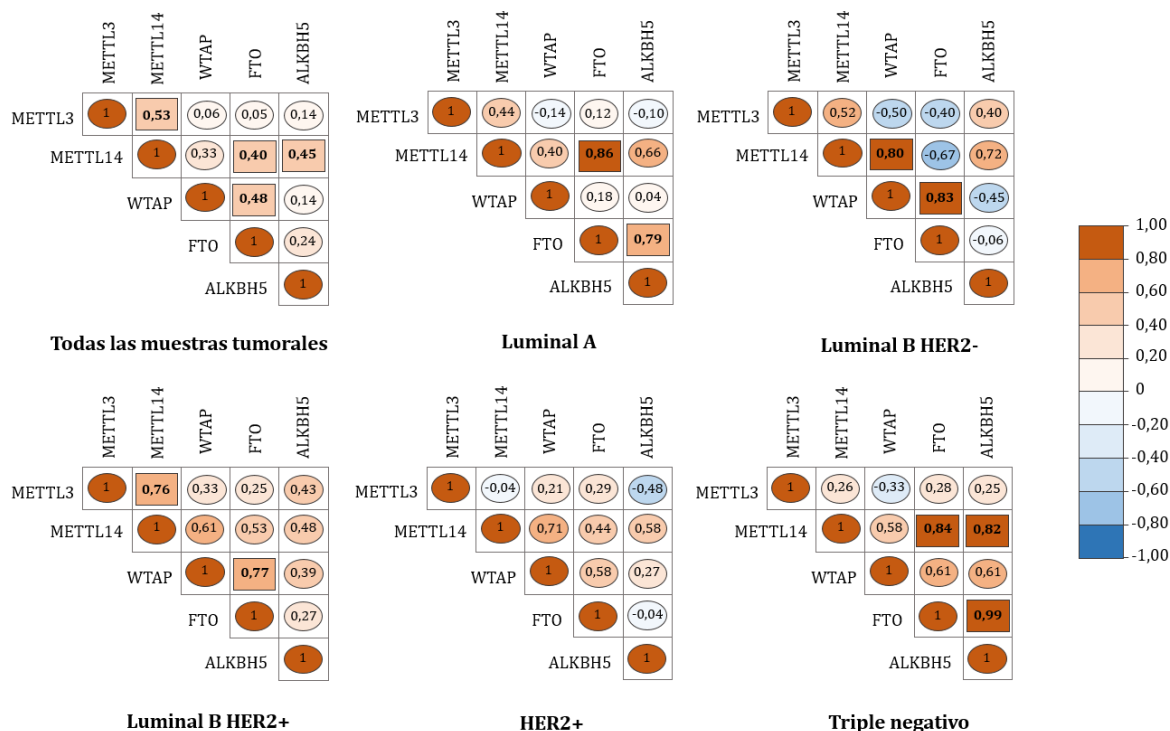


Figura 4.2.10. Análisis de correlación de los niveles de expresión génica de *METTL3*, *METTL14*, *WTAP*, *FTO* y *ALKBH5* en las muestras tumorales de las 35 pacientes con cáncer de mama. Esta prueba también se llevó a cabo teniendo en cuenta los diferentes subtipos moleculares. Las correlaciones significativas están indicadas en negrita y en un recuadro. Los valores de $p < 0,05$ son considerados como significativos.

4.3. IDENTIFICACIÓN *IN SILICO* Y ANÁLISIS DE EXPRESIÓN *IN VITRO* DE m⁶A-SNPs RELACIONADOS CON EL CÁNCER DE MAMA

4.3.1. m⁶A-SNPs relacionados al cáncer de mama

Dado que diversas investigaciones han confirmado que los m⁶A-SNPs podrían influir en distintas enfermedades, se ha llevado a cabo una búsqueda de m⁶A-SNPs que podrían estar implicados en el cáncer de mama mediante el uso de distintas bases de datos. En total, se identificaron 981 m⁶A-SNPs al comparar 69.713 SNPs relacionados con el cáncer de mama (65.331 de GDC, 1.775 del GWAS Catalog, 71 de SNPedia, 1.864 de miRdSNP, 42 de miRNASNP, 43 de MSDD y 587 de SomamiR) con 301.875 m⁶A-SNPs (301.529 m⁶A-SNPs presentes en el ARNm y 346 presentes en el miARN) (Figura 4.3.1). De estos 981 m⁶A-SNPs, 26 presentaban un alto nivel de confianza, 147 mostraban un nivel de confianza medio y 808 un nivel de confianza bajo. Estos diferentes niveles de confianza se relacionan con el método de identificación de los m⁶A-SNPs (145). Los métodos MiCLIP y PA-m⁶A-seq proporcionan datos de alta confianza, mientras que MeRIP-seq proporciona datos de confianza media y la predicción de todo el genoma, basada en el algoritmo Random Forest, proporciona datos con poca confianza.

De los 981 m⁶A-SNPs, todos menos uno, se encuentran en el ARNm. Este último m⁶A-SNP se encuentra en un miARN. Está descrito que la mayoría de los sitios m⁶A metilados se encuentran en la región 5'-UTR, en exones largos, cerca de los codones de terminación, en la región 3'-UTR y en los ARNnc (33,36–38). En nuestra búsqueda, hemos hallado que la mayoría de los m⁶A-SNPs, es decir, los SNPs ubicados dentro o cerca de los motivos m⁶A, se distribuyen en la secuencia codificante (90,62%), mientras que el 6,63% se encuentran en la región 3'-UTR, el 2,65% en la región 5'-UTR y solo el 0,10% en el miARN.

4.3.2. Análisis *cis*-eQTL

Para explorar el posible efecto de los 981 m⁶A-SNPs en el cáncer de mama, investigamos si presentaban un efecto *cis*-eQTL. Teniendo en cuenta los datos del GTEx Portal para muestras de tejido tumoral mamario, ninguno de estos m⁶A-SNPs presentaron un efecto *cis*-eQTL. Sin embargo, cuando realizamos la búsqueda en la base de datos HaploReg, cuatro m⁶A-SNP (rs1801270, rs9379084, rs76563149 y rs11614913) mostraron un efecto *cis*-eQTL en diferentes células y/o tejidos (Figura 4.3.1).

El m⁶A-SNP rs1801270 (C/A) presentó un efecto *cis*-eQTL en fibroblastos (187). Este m⁶A-SNP es una variante sin sentido localizada en el gen inhibidor de quinasa dependiente de ciclina 1A (*CDKN1A*, ID de gen: 1026) en el cromosoma 6 (Tabla 4.3.1). La variante A de este m⁶A-SNP se encuentra en un motivo DRACH incompleto (AGA* CG, donde A* es la adenosina metilada). Además, este m⁶A-SNP está cerca de un motivo DRACH completo y uno incompleto (a 16 nts secuencia arriba y a cinco nts secuencia abajo en el ARN, respectivamente). La frecuencia del alelo A de este m⁶A-SNP se obtuvo de la base de datos Ensembl Genome Browser con el fin de valorar cual podría ser su incidencia. Se observó que la frecuencia variaba según la población, entre 0,067 y 0,488 para las poblaciones europeas (EUR) y del este asiático (EAS), respectivamente (Tabla 4.3.2). Por otro lado, de acuerdo con los conjuntos de datos del factor de transcripción ENCODE CHIP-seq datasets (194) obtenido en la base de datos HaploReg, este m⁶A-SNP podría estar involucrado en la regulación de la expresión génica al alterar el motivo regulador, es decir, el sitio de unión de factores de transcripción y al modificar el sitio de unión de las proteínas POL2 y ZBTB33.

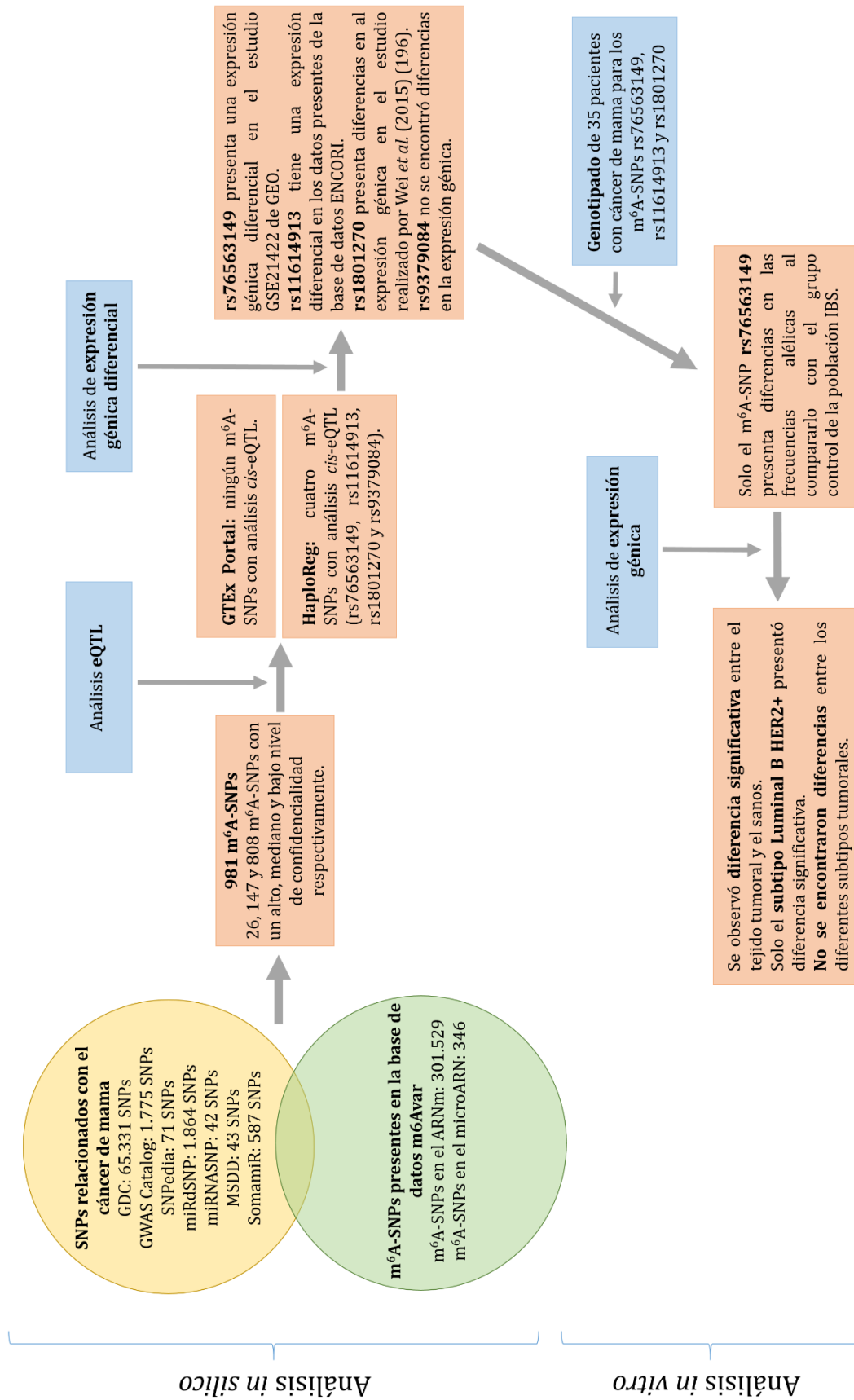


Figura 4.3.1. Diagrama de flujo del diseño del estudio y sus principales resultados. IBS: población Ibérica en España.

Otro m⁶A-SNP es rs9379084 (G/A) que presentó un efecto *cis*-eQTL en sangre total. Este m⁶A-SNP es una variante sin sentido que se encuentra en el gen de la proteína de unión a elemento sensible a Ras 1 (*RREB1*, ID de gen: 6239) en el cromosoma 6 (Tabla 4.3.1). Este m⁶A-SNP forma parte de un motivo DRACH incompleto (TGA*AC). Cerca de este m⁶A-SNP se encuentra un motivo DRACH completo ubicado a siete nts secuencia abajo y dos motivos DRACH incompletos a 14 y 25 nts secuencia arriba en el ARN. La Tabla 4.3.2 muestra la frecuencia del alelo A en las diferentes poblaciones, siendo la frecuencia máxima de 0,159 en EAS y una frecuencia mínima de 0,006 en poblaciones africanas (AFR). En la población EUR, la frecuencia del alelo A es de 0,132. Por otra parte, este m⁶A-SNP altera un motivo VRD, lo que podría resultar en una regulación alterada de la expresión génica (Tabla 4.3.1).

El m⁶A-SNP rs76563149 (G/T) mostró un efecto *cis*-eQTL en diferentes células y tejidos: fibroblasto, tejido adiposo, arteria, esófago, pulmón, músculo esquelético, nervio tibial, estómago, testículo, tiroides y sangre entera. Este m⁶A-SNP aparece en la región 5'-UTR del gen de la proteína de dedo de zinc 354A (*ZNF354A*, ID de gen: 6940) en el cromosoma 5 (Tabla 4.3.1). Dado que este gen está orientado en la hebra negativa, las variantes de este m⁶A-SNP en el ARNm serán C/A. La variante A de este m⁶A-SNP pertenece a un motivo DRACH completo (GAA*GC). En este caso, la frecuencia del alelo A se encontraba en el rango de 0,081 a 0,368 para las poblaciones AFR y del sur asiático (SAS), respectivamente (Tabla 4.3.2). En la población EUR, se observó una frecuencia del alelo A de 0,201. Además, este m⁶A-SNP podría estar involucrado en la regulación de la expresión génica alterando un motivo regulador y la unión de 11 proteínas.

Tabla 4.3.1. m⁶A-SNPs identificados que presentaron señal *cis*-eQTL.

SNP	Posición ^a	Gen	Tipo de mutación	Cadena	Unión de proteínas	Cambio de motivo
rs1801270	6:36684194	<i>CDKN1A</i>	Sin sentido	+	POL2 y ZBTB33	Ascl2, ELF1 HEN1 y Sin3AK-20
rs9379084	6:7231610	<i>RREB1</i>	Sin sentido	+	-	VDR
rs76563149	5:178730572	<i>ZNF354A</i>	5'-UTR	-	POL2, SIN3AK20, POL24H8, TAF1, E2F1, HAE2F1, TR4, CCNT2, HEY1, HMGN3, y ZNF263	EBF y Ets
rs11614913	12:53991815	<i>MIR196A2</i>	ARNnc	+	-	HMG-IY

^a Versión del genoma: GRCh38.p13

Finalmente, el m⁶A-SNP rs11614913 (C/T) mostró efecto *cis*-eQTL en diferentes células y tejidos: fibroblasto, arteria y nervio tibial, piel y sangre entera. El rs11614913 está presente en el gen miARN 196a-2 (*MIR196A2*, gen ID: 406973) en el cromosoma 12 (Tabla 4.3.1). Este m⁶A-SNP está cerca de dos motivos DRACH completos (a cuatro y 19 nts secuencia arriba en el ARN) y un motivo DRACH incompleto (a 11 nts secuencia arriba en el ARN). En este caso, la frecuencia del alelo C está en el rango de 0,458 a 0,860 para las poblaciones de EAS y AFR, respectivamente, siendo la frecuencia de la población EUR de 0,589 (Tabla 4.3.2). El cambio de la variante U a C no altera la estructura secundaria de este miARN (Figura 4.3.2). Este m⁶A-SNP también podría estar involucrado en la regulación de la expresión génica al alterar un motivo regulador.

4.3.3. Análisis *in silico* de la expresión génica

El análisis *in silico* de la expresión génica se llevó a cabo para los cuatro genes en los que los m⁶A-SNPs mostraron un efecto *cis*-eQTL. Para los m⁶A-SNPs rs1801270, rs9379084 y rs76563149 que están presentes en los genes *CDKN1A*, *RREB1* y *ZNF354A* respectivamente, los niveles de expresión génica se obtuvieron a partir de dos conjuntos de datos de la base de datos GEO. En el caso del m⁶A-SNP rs11614913 presente en el gen *MIR196A2*, el nivel de expresión génica se obtuvo de la base de datos ENCORI. Encontramos que únicamente los genes *ZNF354A*

Tabla 4.3.2. Frecuencias genotípicas y alélicas para los cuatro m⁶A-SNPs en diferentes poblaciones. AFR: africana; AMR: americana; EAS: este asiático; EUR: europea; y SAS: sur asiático.

m ⁶ A-SNP	Genotipos / Alelos	Frecuencia alélicas y genotípicas				
		AFR	AMR	EAS	EUR	SAS
rs1801270	AA + AC	0,526	0,421	0,716	0,127	0,262
	CC	0,474	0,579	0,284	0,873	0,738
	A	0,309	0,249	0,488	0,067	0,145
	C	0,691	0,751	0,512	0,933	0,885
rs9379084	AA + AG	0,012	0,216	0,284	0,247	0,202
	GG	0,988	0,784	0,716	0,753	0,798
	A	0,006	0,111	0,159	0,132	0,106
	G	0,994	0,889	0,841	0,868	0,894
rs76563149*	TT + GT	0,156	0,412	0,357	0,508	0,593
	GG	0,844	0,588	0,643	0,492	0,407
	T	0,081	0,235	0,201	0,298	0,368
	G	0,919	0,765	0,799	0,702	0,632
rs11614913	CC + TC	0,977	0,838	0,714	0,817	0,930
	TT	0,023	0,162	0,286	0,183	0,070
	C	0,860	0,611	0,458	0,589	0,742
	T	0,140	0,389	0,542	0,411	0,258

*Las variantes para este m⁶A-SNP en el ARNm son A o C.

y *MIR196A2* presentaron una expresión génica diferencial. Con respecto a *ZNF354A*, la expresión diferencial se obtuvo en el estudio GSE21422 ($p = 0,019$) de cinco muestras de tejido sano y 14 muestras de carcinoma (nueve para carcinomas ductales *in situ* y cinco para carcinomas ductales invasivos). En el caso de *MIR196A2*, se encontró una diferencia significativa en la expresión génica cuando se compararon 1.085 muestras tumorales de carcinoma invasivo de mama con 104 tejidos normales ($p = 4,0e-14$).

Aunque *CDKN1A* no presentó diferencias significativas en la expresión génica en los dos conjuntos de datos analizados, el estudio llevado a cabo por Wei *et al.* (2015) sugirió lo contrario (196). Wei *et al.* analizaron la expresión génica y proteica de *CDKN1A* en 65 muestras tumorales de tejido mamario y sus correspondientes tejidos adyacentes no cancerosos, mostrando que las diferencias en la expresión eran estadísticamente significativas ($p < 0,01$). *RREB1* no mostró diferencias estadísticamente significativas después de analizar los datos de la base de datos GEO. Del mismo modo, no se encontró ningún estudio que analizase la expresión génica de *RREB1* en el tejido mamario.

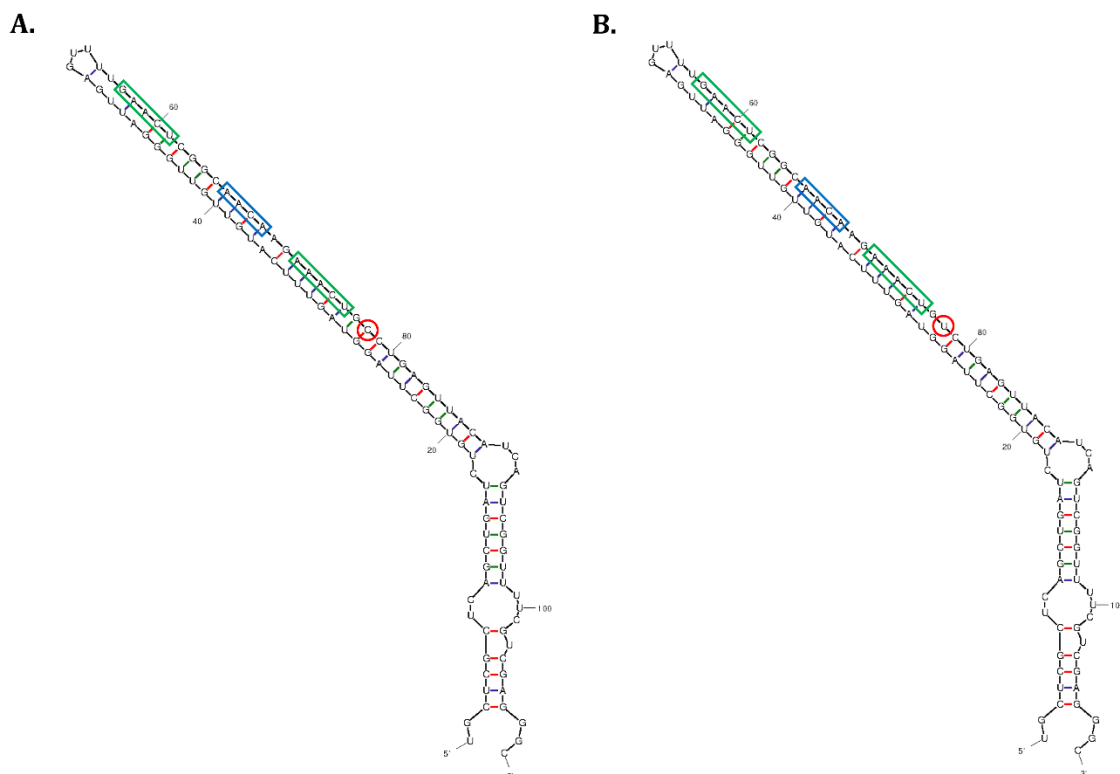


Figura 4.3.2. Representación gráfica de la estructura del miARN miR-196A obtenida de la página web UNAFold Web Server (<http://www.unafold.org/mfold/applications/rna-folding-form.php>) (195). El m⁶A-SNP rs11614913 (C/U) se encuentra dentro del círculo rojo y puede influir en la modificación m⁶A del motivo DRACH completo (recuadros verdes) e incompleto (recuadro azul). **A-** Representación de la estructura del miARN con la variante C para el m⁶A-SNP rs11614913. **B-** Representación de la estructura del miARN con la variante U para el m⁶A-SNP rs11614913.

4.3.4. Validación de los m⁶A-SNPs en muestras de cáncer de mama

El genotipo de las 35 pacientes con cáncer de mama fue determinado mediante secuenciación Sanger para los tres m⁶A-SNPs que presentaron diferencias en el análisis de expresión génica *in silico* (rs1801270, rs76563149 y rs11614913), con el fin de calcular las frecuencias alélicas y genotípicas. Estos resultados han sido comparados con los datos disponibles en la base de datos *1000 Genomes* para mujeres pertenecientes a la población IBS. En el grupo de pacientes con cáncer de mama, el porcentaje de individuos que presentaban el alelo A fue de 11,40% (AA = 5,20% y AC = 5,20%) para el m⁶A-SNP rs1801270 y 65,70% (AA = 14,30% y AC = 51,40%) para el m⁶A-SNP rs76563149 (Tabla 4.3.3). En el caso del m⁶A-SNP rs11614913, el porcentaje de los individuos que presentan el alelo C, que es la variante asociada al cáncer de mama (197–201), fue de 80,00% (CC = 51,40% y CT = 28,60%). Por otro lado, en la población IBS, el 17,00% (AA = 0% y AC = 17,00%) y el 45,30% (AA = 3,80% y AC = 41,50%) presentaron el alelo A para los m⁶A-SNPs rs1801270 y rs76563149, respectivamente, mientras que un 83,00% (CC = 43,40% y CT = 39,60%) presentó el alelo C para el m⁶A-SNP rs11614913.

No se detectaron diferencias estadísticamente significativas entre las pacientes con cáncer de mama y el grupo de mujeres de la población IBS cuando se compararon las frecuencias genotípicas para los tres m⁶A-SNPs (Tabla 4.3.3). Sin embargo, al compararse las frecuencias alélicas, se observó diferencias estadísticamente significativas en el m⁶A-SNP rs76563149 ($p = 0,029$).

Tabla 4.3.3. Frecuencias genotípicas, alélicas y el valor de la prueba estadística χ^2 para los pacientes con cáncer de mama y los controles de la población IBS.

m ⁶ A-SNP/Gen	Alelos / Genotipos	Frecuencias		χ^2	Valor de p
		Muestras Control (n = 53)	Muestras Tumorales (n = 35)		
rs1801270 <i>CDKN1A</i>	AA + AC	0,170	0,114	0,516	0,472
	CC	0,830	0,886		
rs76563149* <i>ZNF354A</i>	A	0,850	0,860	<0,001	0,985
	C	0,915	0,914		
	TT + GT	0,453	0,657	3,536	
	GG	0,547	0,343		
rs11614913 <i>MIR196A2</i>	T	0,245	0,400	4,745	0,029
	G	0,755	0,600		
	CC + CT	0,830	0,800	0,129	
	TT	0,170	0,200		
	C	0,642	0,657	0,045	0,832
	T	0,358	0,343		

*Las variantes para este m⁶A-SNP en el ARNm son A o C.

4.3.5. Análisis de expresión génica para las posiciones m⁶A-SNPs

Dado que el m⁶A-SNP rs76563149 localizado en el gen *ZNF354A* fue el único que presentó diferencia significativa en la frecuencia del alelo A entre pacientes y población IBS, se procedió a evaluar si había diferencias en cuanto a la expresión génica entre el tejido tumoral y el tejido sano mediante RT-qPCR. Para ello, primeramente, se determinó la eficiencia de la pareja de cebadores diseñados para amplificar dicho m⁶A-SNP obteniendo un valor de 90,90% (Figura 4.3.3). La eficiencia de los cebadores que amplifican los genes de referencia se determinó previamente en el apartado 4.2.2.

Al comparar la expresión, se observó que el tejido tumoral y el tejido sano presentaban diferencias significativas ($p < 0,001$), encontrándose una expresión reducida en las muestras tumorales ($0,785 \pm 0,613$ en las muestras tumorales y $1,095 \pm 0,564$ en las muestras sanas). Posteriormente, al evaluar la expresión de cada subtipo molecular con su correspondiente tejido sano, únicamente el subtipo luminal B HER2+ presentó diferencia significativa ($p = 0,001$), con una expresión reducida en las muestras tumorales (Figura 4.3.4A-E y Tabla 4.3.4). No se observaron diferencias significativas al comparar la expresión entre los subtipos tumorales ($p = 0,117$) (Figura 4.3.4F y Tabla 4.3.4). Por lo que, este m⁶A-SNP podría ser de gran interés como marcador epitranscriptómico del cáncer de mama, y, específicamente del subtipo luminal B HER2+.

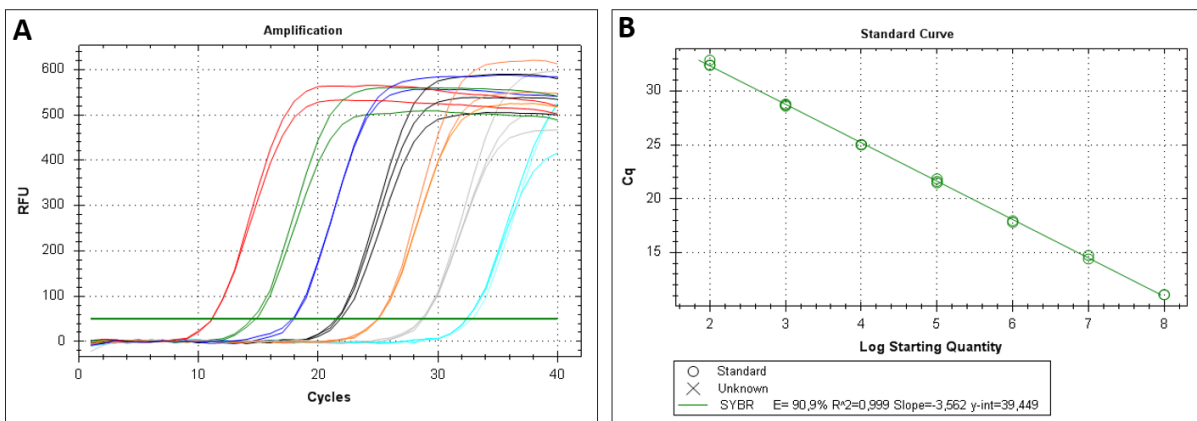


Figura 4.3.3. Análisis de eficiencia de los cebadores que amplifican el m⁶A-SNP rs76563149 presente en el gen *ZNF354A*. **A-** Curvas de amplificación de las diluciones realizadas con diferentes números de copias. **B-** Recta obtenida a partir de los valores Cq de las curvas de amplificación. E: eficiencia; R²: coeficiente de correlación; Slope: pendiente; Y-int: valor de intersección con el eje y. Concentraciones de las diluciones: Rojo: 1x10⁸ copias/μL; Verde: 1x10⁷ copias/μL; Azul: 1x10⁶ copias/μL; Negro: 1x10⁵ copias/μL; Naranja: 1x10⁴ copias/μL; Gris: 1x10³ copias/μL; Celeste: 1x10² copias/μL.

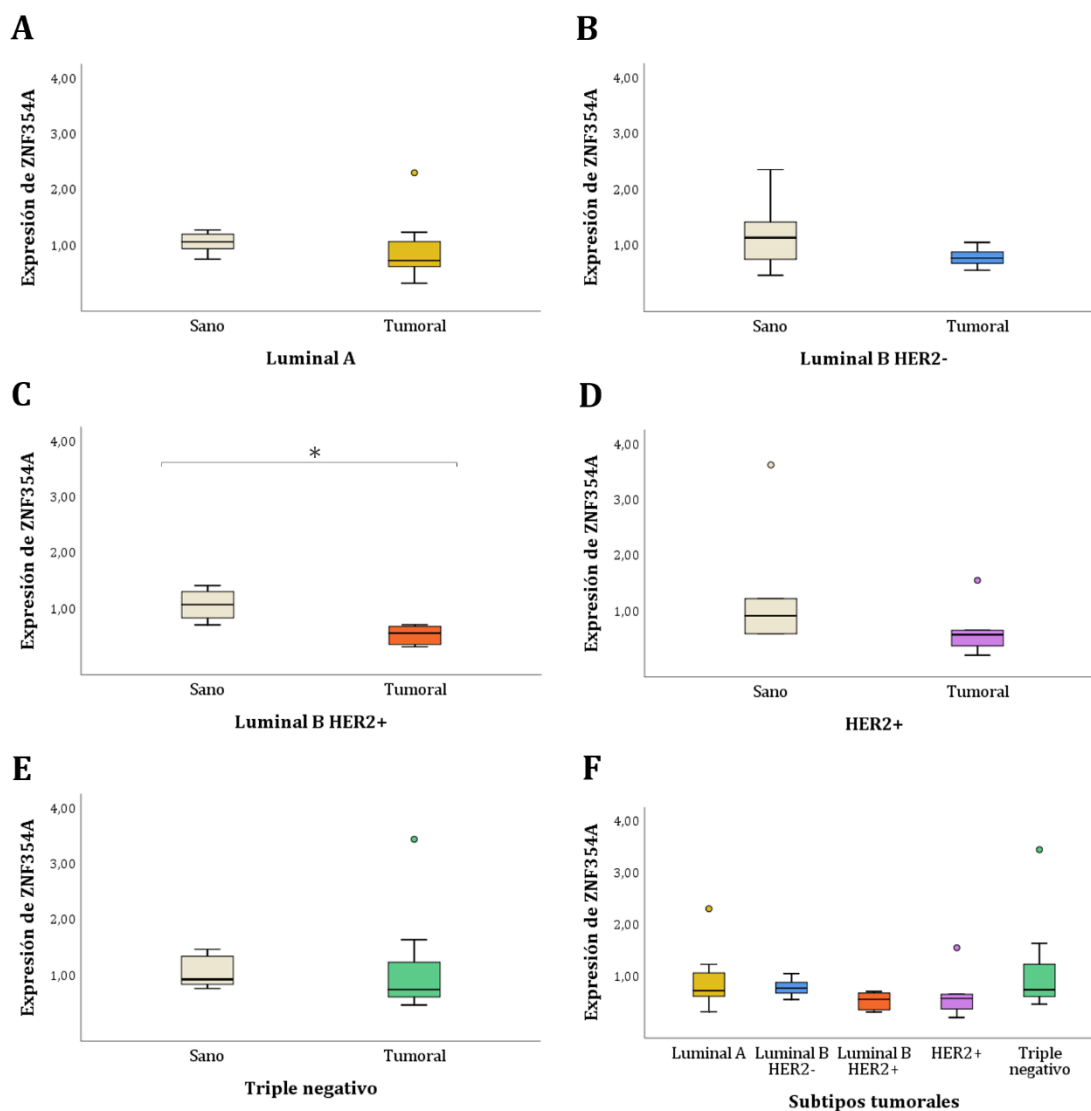


Figura 4.3.4. Niveles de expresión relativa del gen *ZNF354A*. El fragmento analizado contenía el m⁶A-SNP rs76563149. Expresión génica entre cada subtipo tumoral y sus correspondientes tejidos sanos: **A-** Subtipo luminal A, **B-** Subtipo luminal B HER2-, **C-** Subtipo luminal B HER2+, **D-** Subtipo HER2+ y **E-** Subtipo triple negativo. **F-** Expresión génica entre los diferentes subtipos tumorales. * *P* <0,05. Los círculos son valores atípicos.

Tabla 4.3.4. Promedio de la expresión génica relativa para el m⁶A-SNP rs76563149 localizado en el gen *ZNF354A* con su desviación estándar (DS) en el tejido tumoral y su correspondiente tejido sano y los valores *p* obtenidos para las pruebas U de Mann-Whitney y H de Kruskal-Wallis.

Pruebas estadísticas		Luminal A	Luminal B HER2-	Luminal B HER2+	HER2+	Triple negativo
Promedio ± DS	Sano	1,016 ± 0,191	1,151 ± 0,645	1,032 ± 0,270	1,274 ± 1,161	1,037 ± 0,304
	Tumoral	0,924 ± 0,657	0,752 ± 0,167	0,492 ± 0,163	0,622 ± 0,468	1,156 ± 1,064
U de Mann-Whitney (<i>p</i>)		0,180	0,180	0,001	0,109	0,277
H de Kruskal-Wallis (<i>p</i>)		0,117				

4.4. DETERMINACIÓN DE POSICIONES m⁶A METILADAS

4.4.1. Posiciones m⁶A metiladas en el cáncer de mama

El análisis mediante MeRIP-seq se llevó a cabo en 20 muestras de tejido tumoral pertenecientes a los diferentes subtipos moleculares de cáncer de mama: luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo (cuatro muestras por cada subtipo) con el fin de determinar el perfil de metilación de cada subtipo tumoral. En total, se detectaron 1.703 picos m⁶A. El subtipo luminal B HER2- presentó el mayor número de picos m⁶A con 873 picos, mientras que el subtipo HER2+ presentó el menor número con 55 picos. En el caso de los subtipos luminal A, luminal B HER2+ y triple negativo, el número de picos m⁶A identificados fue de 339, 253 y 183, respectivamente. Los picos m⁶A encontrados podrían ser específicos de cada subtipo molecular o podrían ser compartidos, es decir, podrían encontrarse en dos o más subtipos. El número de picos m⁶A específicos para cada subtipo fue de: 241, 734, 168, 45 y 112 para luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo, respectivamente (Figura 4.4.1A).

Posteriormente, evaluamos el número de genes cuyos transcriptos presentaban picos m⁶A. Encontramos que los picos m⁶A presentes en el subtipo luminal A se correspondían con transcriptos de 312 genes, mientras que en el subtipo luminal B HER2- se correspondían con transcriptos de 783 genes, en el subtipo luminal B HER2+ con 237 genes, en el subtipo HER2+ con 51 genes y en el subtipo triple negativo con transcriptos de 177 genes. Al analizar el número de picos m⁶A por transcripto (gen), observamos que la gran mayoría tenía un pico m⁶A (Figura 4.4.1B). Solo dos subtipos presentaron transcriptos con tres picos m⁶A (en tres genes para el subtipo luminal A y en cuatro genes para el subtipo luminal B HER2-) y, únicamente, el subtipo luminal B HER2- presentó cuatro picos m⁶A en el transcripto de un solo gen.

Por otra parte, se evaluó la distribución de los picos m⁶A en todo el transcriptoma para cada subtipo molecular. Las posiciones de los picos m⁶A se clasificaron principalmente según que estuvieran en exones, intrones, región 3' y 5'-UTR, promotor, sitio de inicio de la transcripción (TSS), región intergénica y ARNnc. Como se muestra en la Figura 4.4.1C, en todos los subtipos, los exones se encontraron especialmente enriquecidos en picos m⁶A (45,45-58,47%), así como la región 3'-UTR (18,58-25,45%). Por otro lado, el subtipo luminal A presentó un enriquecimiento de picos m⁶A en el promotor (4,42%) en comparación con el resto de subtipos

moleculares, mientras que los subtipos luminal B HER2- y luminal B HER2+ presentaron una disminución de picos m⁶A en los intrones (2,75%) y en TSS (2,37%), respectivamente. El

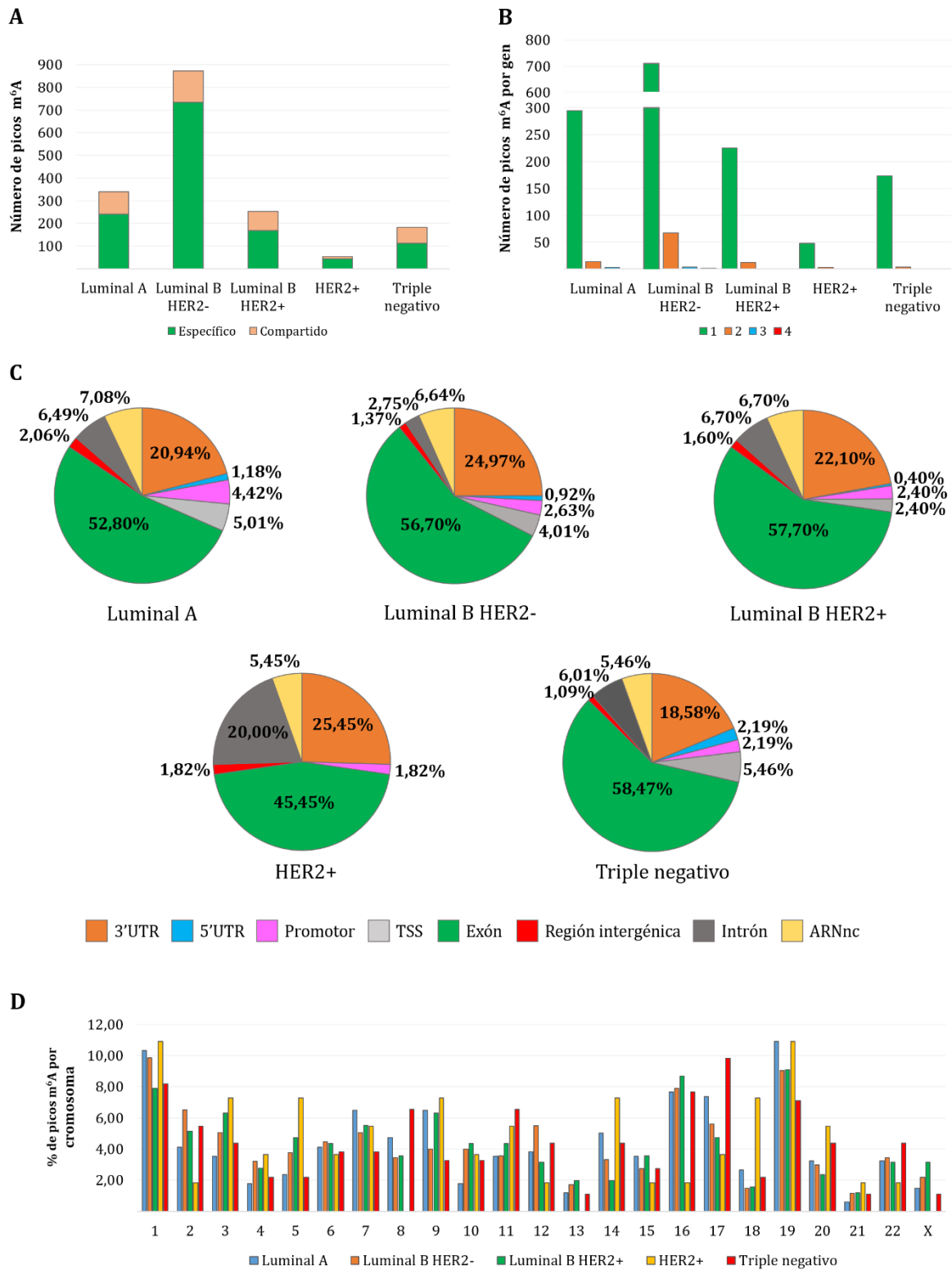


Figura 4.4.1. Análisis de los picos m⁶A en los distintos subtipos moleculares. **A-** Número de picos m⁶A, específicos y compartidos. **B-** Número de picos m⁶A por gen. **C-** Distribución de los picos m⁶A en todo el transcriptoma. **D-** Distribución de los picos m⁶A en los cromosomas.

subtipo HER2+ no presentó picos m⁶A en la región 5'-UTR y en TSS, y en comparación con el resto de subtipos, presentó un enriquecimiento de picos m⁶A en los intrones (20,00%). También, se analizó la distribución de los picos m⁶A por cromosoma. La Figura 4.4.1D muestra el porcentaje de picos m⁶A por cromosoma para cada subtipo molecular. Como se puede observar, para todos los subtipos, el mayor porcentaje de picos se encontró en los cromosomas 1 y 19. El subtipo HER2+ no presentó picos m⁶A en los cromosomas 8, 13 y X.

Los 20 picos m⁶A con mayor *peak score* de cada subtipo molecular se enumeran en la Tabla 4.4.1. Los picos m⁶A ubicados en la última posición de cada subtipo molecular presentaron valores de *p* de 1,74e-33 para el subtipo luminal A, 2,24e-130 para el luminal B HER2-, 7,49e-31 para el luminal HER2+, 2,95e-16 para el subtipo HER2+ y 9,33e-31 para el subtipo triple negativo. Por lo que, estos valores de *p* muestran la elevada fiabilidad de los picos m⁶A encontrados. Como se puede observar, en todos los subtipos, la mayoría de los picos m⁶A se encuentran en los exones. En el subtipo HER2+, de los 20 picos m⁶A, 19 son específicos del subtipo, mientras que, en el subtipo triple negativo, únicamente tres picos m⁶A son específicos y el resto son compartidos, es decir, se encuentran en dos o más subtipos. También, se puede observar que el gen epipaklin 1 (*EPPK1*) se encuentra presente en todos los subtipos moleculares, a excepción de HER2+.

Tabla 4.4.1. Top 20 de los picos m⁶A encontrados en los distintos subtipos moleculares. Los picos m⁶A se han ordenado en función del *peak score* (valor de *p* transformado a $-\log_{10}(p)$). CHR: cromosomas.

Subtipo molecular	Peak Score	CHR	Inicio	Fin	Gen	Específico / compartido	Distribución
Luminal A	220,06	17	79833890	79834672	<i>CBX4</i>	Compartido	Exón
	107,66	8	143865034	143865809	<i>EPPK1</i>	Compartido	Exón
	106,21	8	143860095	143860743	<i>EPPK1</i>	Compartido	Exón
	71,41	8	144850435	144851308	<i>COMMD5</i>	Compartido	Exón
	51,49	14	105250371	105250784	<i>BTBD6</i>	Compartido	Exón
	51,45	2	96251827	96253034	<i>TMEM127</i>	Compartido	3'UTR
	50,30	16	30739028	30739299	<i>SRCAP</i>	Específico	Exón
	46,34	17	81244977	81246621	<i>SLC38A10</i>	Compartido	Exón
	44,12	7	40094432	40095103	<i>CDK13</i>	Compartido	Exón
	44,02	15	90881306	90882263	<i>FURIN</i>	Compartido	Exón
	41,49	19	42198741	42199748	<i>DEDD2</i>	Compartido	ARNnc
	40,20	16	31080089	31080604	<i>ZNF646</i>	Específico	Exón
	39,46	19	45712033	45712716	<i>FBXO46</i>	Compartido	Exón
	34,90	17	79833660	79833864	<i>CBX4</i>	Específico	3'UTR

Subtipo molecular	Peak Score	CHR	Inicio	Fin	Gen	Específico / compartido	Distribución
Luminal A	34,74	1	109928940	109929611	<i>CSF1</i>	Específico	3'UTR
	34,68	22	17109061	17109893	<i>IL17RA</i>	Compartido	Exón
	34,32	1	27548470	27549363	<i>AHDC1</i>	Específico	Exón
	34,10	4	1250488	1250914	<i>CTBP1</i>	Específico	ARNnc
	32,88	7	105112546	105112800	<i>KMT2E</i>	Específico	Exón
	32,76	19	48452711	48453395	<i>GRWD1</i>	Compartido	3'UTR
Luminal B HER2-	540,58	2	216674394	216674923	<i>IGFBP5</i>	Específico	3'UTR
	434,11	12	76030415	76030963	<i>PHLDA1</i>	Específico	Exón
	346,05	3	49531258	49533435	<i>DAG1</i>	Específico	Exón
	269,30	8	143863429	143863856	<i>EPPK1</i>	Específico	Exón
	249,18	8	41933010	41933527	<i>KAT6A</i>	Específico	Exón
	243,13	8	95154385	95155002	<i>PLEKHF2</i>	Específico	Exón
	242,20	22	24542246	24543107	<i>GUCD1</i>	Específico	ARNnc
	220,06	17	79833890	79834672	<i>CBX4</i>	Compartido	Exón
	201,32	10	79312641	79313385	<i>ZMIZ1</i>	Específico	3'UTR
	201,02	22	41176861	41178082	<i>EP300</i>	Específico	Exón
	199,50	3	9475440	9476160	<i>SETD5</i>	Específico	Exón
	196,98	1	155016218	155016906	<i>ZBTB7B</i>	Específico	Exón
	183,32	8	41932340	41933004	<i>KAT6A</i>	Específico	Exón
	176,68	8	125436028	125436980	<i>TR1B1</i>	Específico	3'UTR
	161,70	X	153954332	153955029	<i>HCFC1</i>	Específico	Exón
	161,58	2	96253164	96253879	<i>TMEM127</i>	Específico	3'UTR
	159,10	11	118627336	118628318	<i>PHLDB1</i>	Específico	Exón
146,15	19	49908629	49909832	<i>NUP62</i>	Específico	Exón	
134,61	8	63186407	63187315	<i>YTHDF3</i>	Específico	Exón	
129,65	20	50891191	50891482	<i>ADNP</i>	Específico	3'UTR	
Luminal B HER2+	109,26	8	143868891	143872380	<i>EPPK1</i>	Compartido	Exón
	63,83	6	108663594	108664963	<i>FOXO3</i>	Compartido	Exón
	62,86	3	153162799	153163364	<i>RAP2B</i>	Compartido	Exón
	40,03	8	60865358	60865936	<i>CHD7</i>	Compartido	Exón
	39,03	3	65356526	65356960	<i>MAGI1</i>	Específico	Exón
	37,78	7	127581848	127582863	<i>GCC1</i>	Compartido	Exón
	36,64	22	30020051	30020691	<i>MTMR3</i>	Específico	Exón
	35,97	5	40691935	40692433	<i>PTGER4</i>	Específico	Exón
	35,63	8	38210294	38210832	<i>BAG4</i>	Específico	3'UTR
	35,36	5	177135618	177135933	<i>NSD1</i>	Específico	Exón
	34,76	19	52798279	52798816	<i>ZNF28</i>	Específico	ARNnc

Continuación de la Tabla 4.4.1.

Subtipo molecular	Peak Score	CHR	Inicio	Fin	Gen	Específico / compartido	Distribución
Luminal B HER2+	34,73	12	106067174	106067938	<i>NUAK1</i>	Específico	Exón
	34,68	22	17109061	17109893	<i>IL17RA</i>	Compartido	Exón
	34,64	1	248812036	248812384	<i>SH3BP5L</i>	Compartido	Exón
	32,48	14	91602782	91603255	<i>CATSPERB</i>	Específico	Intrón
	32,21	19	11949439	11949944	<i>ZNF700</i>	Compartido	Exón
	31,69	11	117998461	117999295	<i>IL10RA</i>	Específico	ARNnc
	31,20	3	15073993	15074186	<i>RBSN</i>	Específico	Exón
	31,18	15	93024617	93025148	<i>CHD2</i>	Específico	3'UTR
	30,10	1	15939287	15939975	<i>SPEN</i>	Compartido	3'UTR
HER2+	58,04	11	112082622	112083251	<i>NKAPD1</i>	Específico	Exón
	56,27	6	116436171	116437159	<i>DSE</i>	Compartido	Exón
	41,67	14	55043258	55043661	<i>SOCS4</i>	Específico	Exón
	40,30	4	48499155	48499669	<i>FRYL</i>	Específico	3'UTR
	34,36	1	37537558	37538014	<i>SNIP1</i>	Compartido	Exón
	30,84	1	181089147	181089681	<i>IER5</i>	Específico	Exón
	28,47	14	80202530	80203243	<i>DIO2</i>	Compartido	Exón
	25,85	18	80135755	80137473	<i>ADNP2</i>	Compartido	Exón
	21,57	9	132895959	132896709	<i>TSC1</i>	Compartido	Exón
	21,27	19	21809402	21809803	<i>ZNF43</i>	Compartido	Exón
	19,97	3	122740303	122740731	<i>HSPBAP1</i>	Específico	Exón
	19,59	11	58967355	58967803	<i>GLYATL1</i>	Específico	Intrón
	19,15	9	107010869	107011203	<i>ZNF462</i>	Específico	3'UTR
	18,30	15	57708744	57708965	<i>POLR2M</i>	Específico	Exón
	18,16	18	80137736	80138957	<i>ADNP2</i>	Compartido	Exón
	18,01	5	649218	649731	<i>CEP72</i>	Específico	Intrón
	17,95	5	37102917	37103517	-	Específico	Región intergénica
	16,67	10	112816822	112817301	<i>VTI1A</i>	Específico	3'UTR
15,98	2	121650796	121651179	<i>NFK-AS1</i>	Específico	Intrón	
15,53	20	10623331	10623537	<i>SLX4IP</i>	Específico	3'UTR	
Triple negativo	109,26	8	143868891	143872380	<i>EPPK1</i>	Compartido	Exón
	107,66	8	143865034	143865809	<i>EPPK1</i>	Compartido	Exón
	46,34	17	81244977	81246621	<i>SLC38A10</i>	Compartido	Exón
	45,04	11	128969017	128969831	<i>ARHGAP32</i>	Compartido	Exón
	44,33	19	42273211	42274560	<i>CIC</i>	Compartido	Exón
	43,93	17	42990842	42991541	<i>RUNDC1</i>	Compartido	Exón
	42,48	19	48480528	48480947	<i>CYTH2</i>	Específico	3'UTR

Continuación de la Tabla 4.4.1.

Subtipo molecular	Peak Score	CHR	Inicio	Fin	Gen	Específico / compartido	Distribución
Triple negativo	39,00	16	84178416	84179511	<i>TAF1C</i>	Compartido	Exón
	37,34	11	66857630	66858513	<i>LRFN4</i>	Compartido	Exón
	37,33	5	61543620	61544312	<i>ZSWIM6</i>	Compartido	Exón
	34,64	1	248812036	248812384	<i>SH3BP5L</i>	Compartido	Exón
	34,36	1	37537558	37538014	<i>SNIP1</i>	Compartido	Exón
	34,28	17	44397562	44398794	<i>GPATCH8</i>	Compartido	Exón
	33,73	19	40441322	40442027	<i>SERTAD3</i>	Compartido	Exón
	30,92	6	7246373	7246612	<i>RREB1</i>	Específico	Exón
	30,91	19	1576656	1578110	<i>MBD3</i>	Compartido	3'UTR
	30,62	20	52152254	52153417	<i>ZFP64</i>	Compartido	Exón
	30,04	16	627364	629241	<i>RAB40C</i>	Compartido	3'UTR
30,03	19	40800991	40801129	<i>EGLN2</i>	Específico	Exón	

Continuación de la Tabla 4.4.1.

4.4.2. Análisis de ontología génica (GO) y de las rutas metabólicas de KEGG

Para investigar el significado biológico de la metilación m⁶A en los distintos subtipos moleculares de cáncer de mama, realizamos un análisis de GO y de las rutas metabólicas de KEGG mediante la base de datos David Bioinformatics Resources, para las proteínas que codifican los genes presentes en el top 20 de cada subtipo. El análisis de GO clasifica los genes en tres grupos funcionales: proceso biológico (BP), componente celular (CC) y función molecular (MF). Las Figuras 4.4.2-4.4.6 muestran los 10 primeros resultados para BP, CC y MF que presentaban significancia estadística para cada subtipo molecular.

En el subtipo luminal A, los tres principales términos de GO enriquecidos fueron modificación de la peptidilisina, regulación del proceso metabólico primario y regulación del proceso metabólico de las macromoléculas para la categoría BP; mota nuclear, complejo de transferasa y cuerpo nuclear para la categoría CC; y actividad del cofactor de transcripción, unión del factor de transcripción y unión a proteínas para la categoría MF (Figura 4.4.2). En el subtipo luminal B HER2-, encontramos que los tres principales enriquecimientos fueron modificación de la peptidilisina, regulación negativa de la expresión génica y regulación negativa del proceso metabólico de macromoléculas para la categoría BP; parte nuclear, complejo de histona acetiltransferasa y parte de nucleoplasma para la categoría CC; y unión del complejo macromoleculas, unión a la cromatina y actividad histona-acetiltransferasa para la categoría MF (Figura 4.4.3). En el subtipo luminal B HER2+, los enriquecimientos para la categoría BP

fueron: respuesta a estímulos, regulación del proceso celular y regulación de la respuesta al estrés; para la categoría CC: unión estrecha bicelular, unión oclusiva y complejo de unión apical; y para la categoría MF: unión al núcleo del promotor, unión a proteínas y unión a compuestos heterocíclicos (Figura 4.4.4). Para el subtipo HER2+ los tres principales enriquecimientos para la categoría BF fueron proceso biosintético de macromoléculas, celular y de sustancias orgánicas; y para la categoría CC: corteza celular, región citoplasmática y sitio del crecimiento polarizado (Figura 4.4.5). Finalmente, en el subtipo triple negativo encontramos transducción de señales intracelulares y de señales mediadas por pequeñas GTPasas y regulación de procesos celulares para la categoría BF; y lumen nuclear para la categoría CC (Figura 4.4.6). Tanto para el subtipo HER2+ como para el subtipo triple negativo no se encontraron enriquecimientos para la categoría MF.

En cuanto al análisis de las rutas metabólicas de KEGG, únicamente encontramos que las proteínas que codificadas por los genes del subtipo HER2+ se encontraban asociadas a la vía de señalización de la insulina.

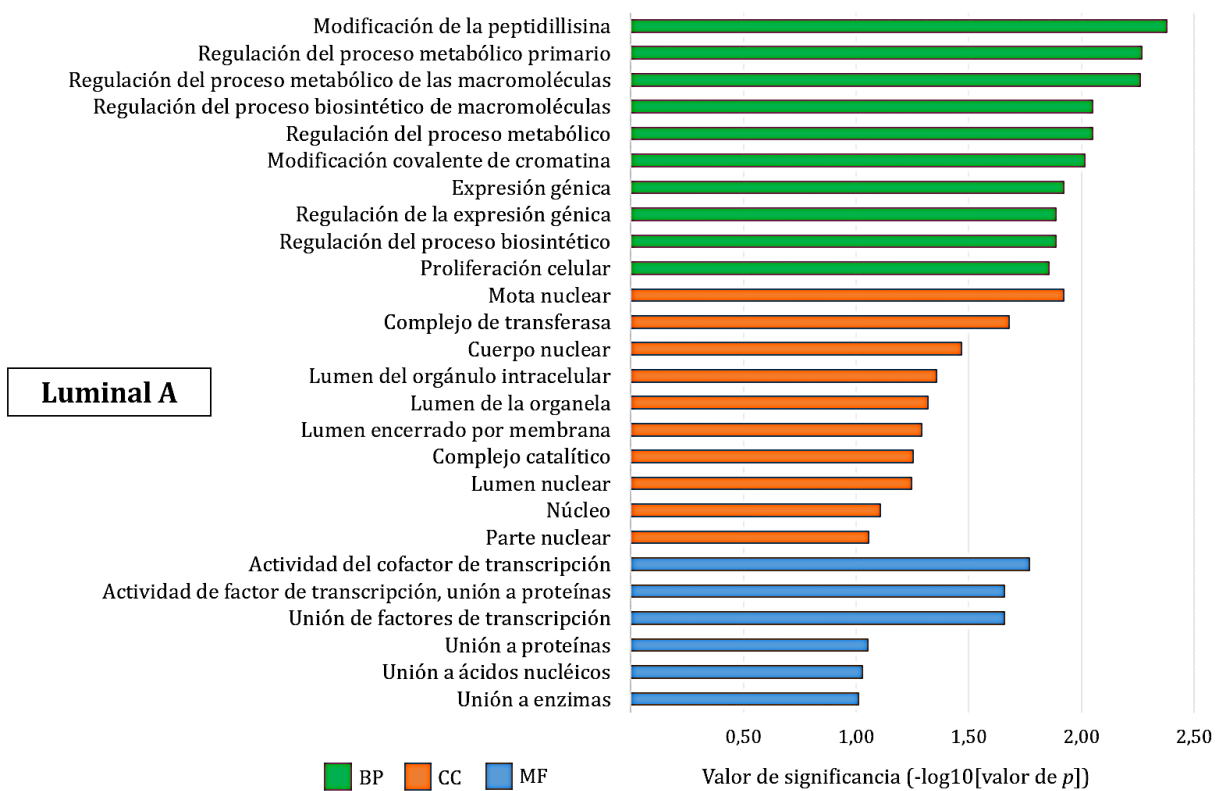


Figura 4.4.2. Lista de los 10 primeros enriquecimientos para las categorías BF, CC y MF obtenidas tras el análisis de GO para el subtipo molecular luminal A.

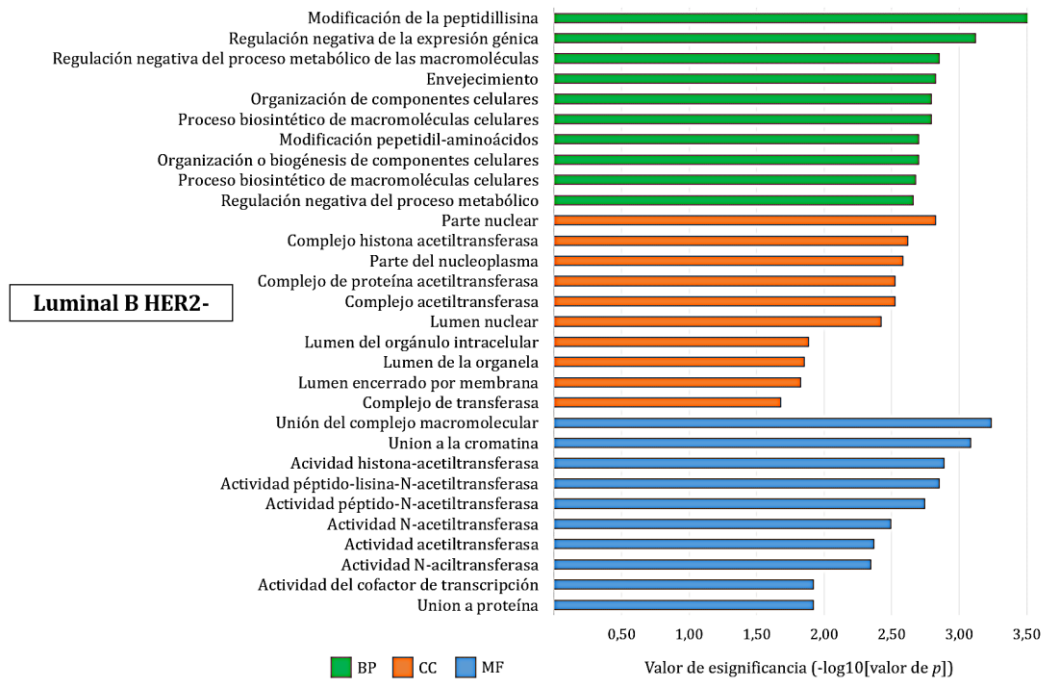


Figura 4.4.3. Lista de los 10 primeros enriquecimientos para las categorías BF, CC y MF obtenidas tras el análisis de GO para el subtipo molecular luminal B HER2-.

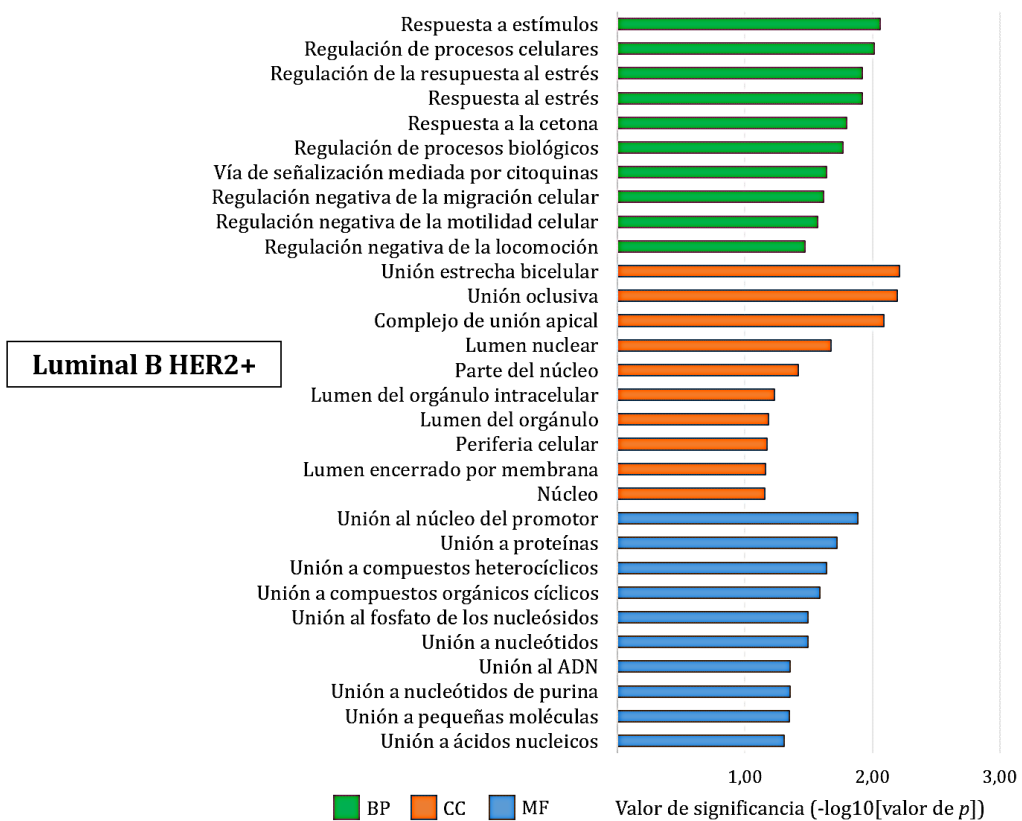


Figura 4.4.4. Lista de los 10 primeros enriquecimientos para las categorías BF, CC y MF obtenidas tras el análisis de GO para el subtipo molecular luminal B HER2+.

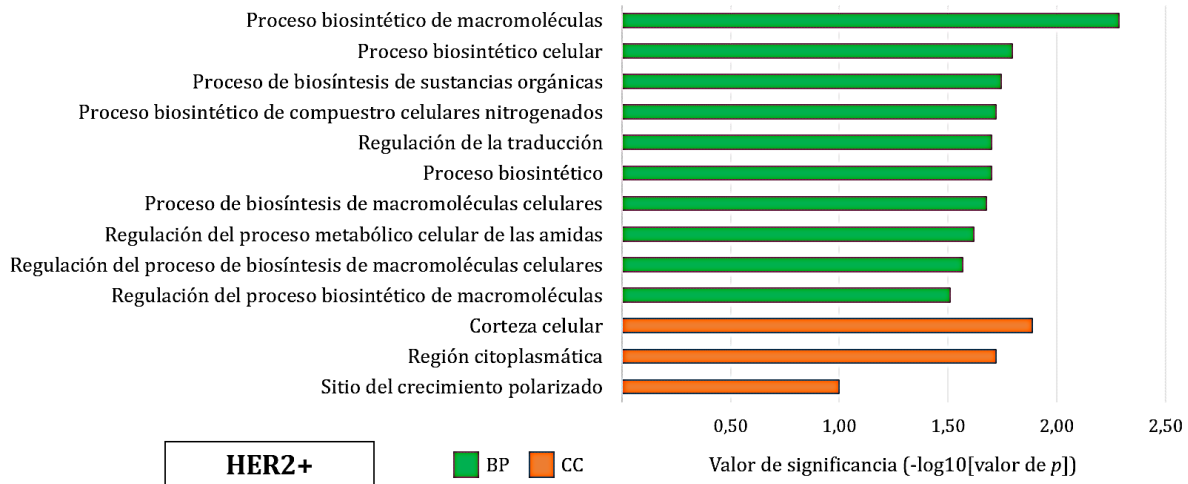


Figura 4.4.5. Lista de los 10 primeros enriquecimientos para las categorías BF, CC y MF obtenidas tras el análisis de GO para el subtipo molecular HER2+.

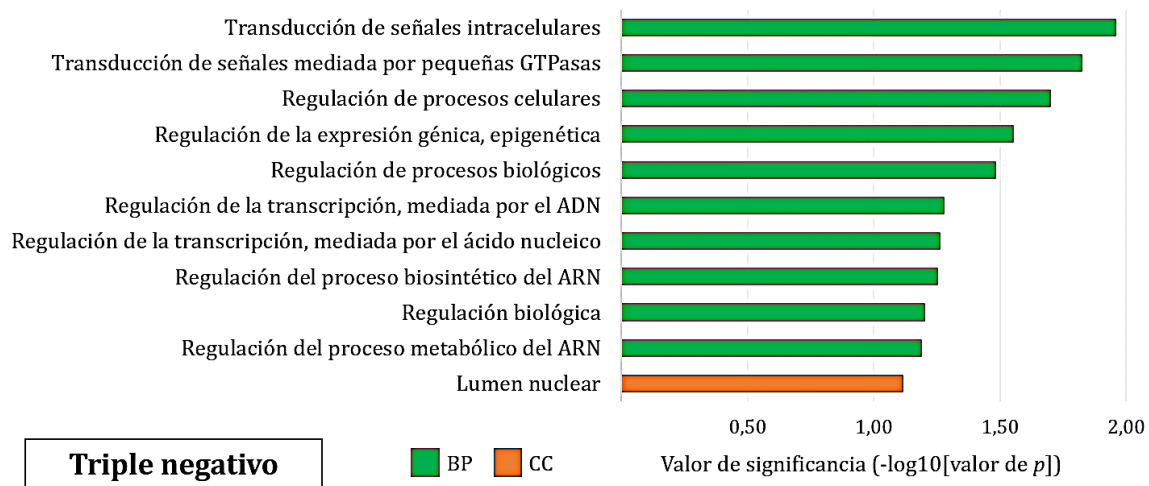


Figura 4.4.6. Lista de los 10 primeros enriquecimientos para las categorías BF, CC y MF obtenidas tras el análisis de GO para el subtipo molecular triple negativo.

4.4.3. Análisis *in silico* de expresión génica

Para determinar si los picos m⁶A encontrados en cada subtipo molecular podrían tener un efecto en la expresión, llevamos a cabo un análisis de expresión génica *in silico* para los cinco primeros genes del top 20 de cada subtipo (Tabla 4.4.1) utilizando los datos de la base de datos UCSC. Para ello se compararon los datos de expresión génica de cada subtipo frente a los datos de expresión del tejido normal. Para los subtipos luminal B HER2- y luminal B HER2+, la base de datos UCSC presenta datos de expresión génica para el subtipo luminal B sin discriminar

entre HER2- o HER2+. Por tal motivo, esto debe de tenerse en cuenta a la hora de interpretar los resultados obtenidos para los subtipos luminal B HER2- y luminal B HER2+.

En la Figura 4.4.7 se pueden observar los resultados obtenidos. En el subtipo luminal A, el gen *EPPK1* presentaba dos picos m⁶A *peak score* elevado, por lo que, dicho gen se encontraba repetido en los primeros cinco genes del top 20. En este subtipo, se pueden observar diferencias significativas en la expresión de los cuatro genes analizados: *CBX4*, *EPPK1*, *COMMD5* y *BTBD6*. Para los genes correspondientes al subtipo luminal B HER2-, se encontraron diferencias significativas en la expresión génica de *IGFBP5*, *PHLDA1* y *EPPK1*. Para el gen *KAT6A*, la base de datos no presentaba información sobre la expresión génica. Por otra parte, se observaron diferencias significativas para cuatro de los genes correspondientes al subtipo luminal B HER2+, es decir, *EPPK1*, *FOXO3*, *CHD7* y *MAG11*. En el caso del subtipo HER2+, de los cinco genes analizados, se observaron diferencias significativas en la expresión de *DSE*, *FRYL* y *SNIP1*. Para el gen *NKAPD1*, no se encontró información sobre la expresión génica. En el subtipo triple negativo, al igual que en el subtipo luminal A, el gen *EPPK1* presentó dos picos m⁶A, por lo que este gen se encontraba repetido. En este caso, únicamente el gen *SLC38A10* no presentó diferencias significativas.

Las posibles funciones de las proteínas codificadas por los genes que presentaron diferencias significativas, se buscaron en la base de datos del NCBI para evaluar la importancia que podrían tener en el proceso tumoral. A continuación, se detallan los resultados obtenidos:

- *CBX4*: codifica una proteína que participa en la regulación negativa de la transcripción mediada por la ARN polimerasa II.
- *EPPK1*: codifica una proteína que reorganiza la arquitectura del citoesqueleto.
- *COMMD5*: la proteína codificada por este gen se prevé que participe en la morfogénesis del túbulo proximal.
- *BTBD6*: la proteína codificada por este gen se prevé que esté involucrada en la neurogénesis.
- *IGFBP5*: el producto de este gen activa la actividad de unión del factor de crecimiento similar a la insulina tipo I.
- *PHLDA1*: la proteína codificada por este gen puede desempeñar un papel importante en los efectos antiapoptóticos del factor de crecimiento similar a la insulina tipo I.

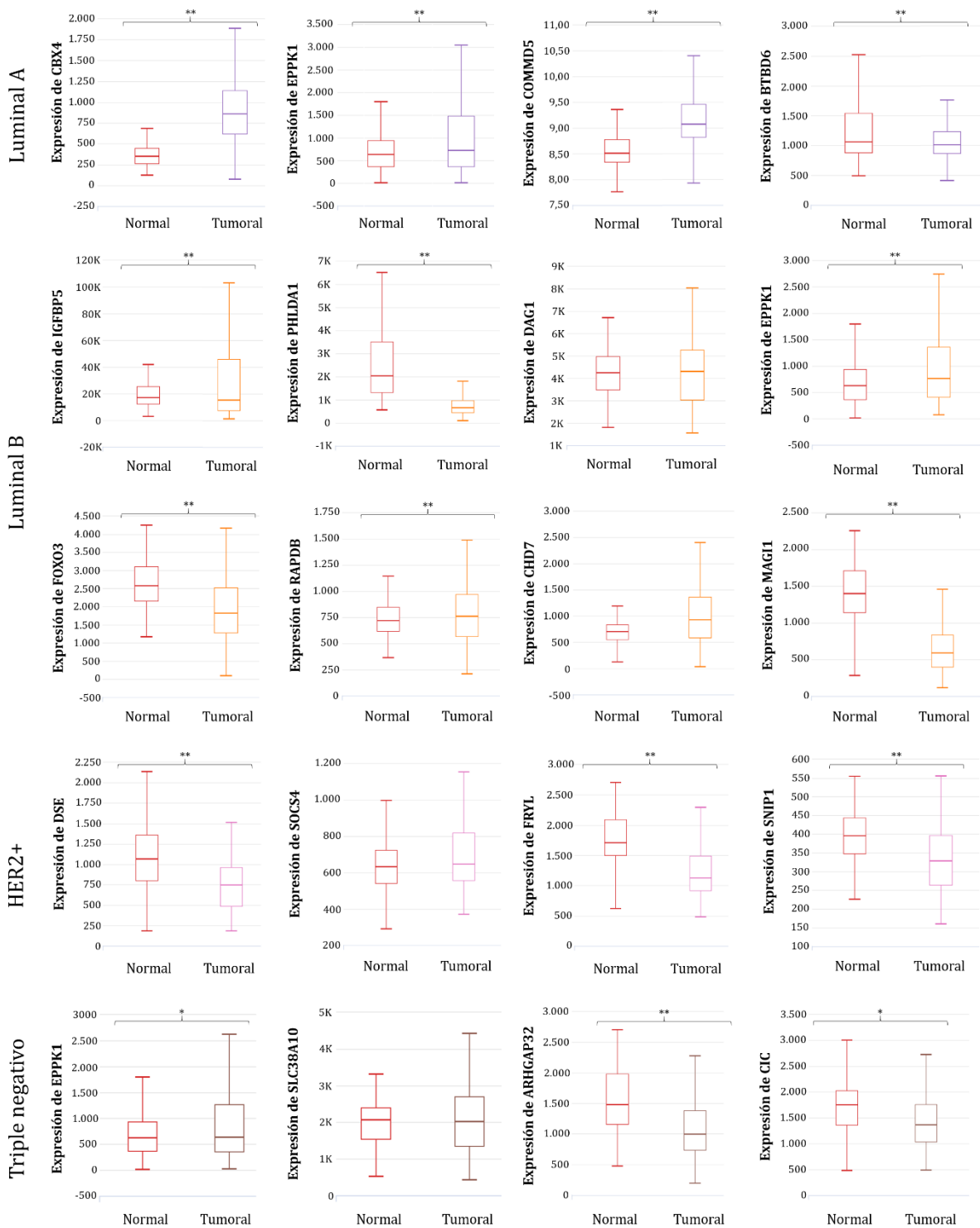


Figura 4.4.7. Análisis de expresión génica *in silico* para los cinco primeros genes del top 20 de cada subtipo molecular. La línea media de cada diagrama de caja representa la mediana. * $P < 0,05$. ** $P < 0,001$.

- *FOXO3*: codifica una proteína que pertenece a una familia de factores de transcripción y es posible que funcione como un desencadenante de la apoptosis.
- *CHD7*: la proteína codificada por este gen probablemente sea un regulador de la transcripción.
- *MAG11*: el producto de este gen puede desempeñar un papel como proteína de andamiaje en las uniones célula-célula.
- *DSE*: el producto de este gen es un antígeno de rechazo tumoral.
- *FRYL*: la proteína codificada por este gen se prevé que participe en la morfogénesis celular y el desarrollo de la proyección neuronal.
- *SNIP1*: la proteína codificada por este gen funciona como un coactivador transcripcional, regula la estabilidad del ARNm de la ciclina D1 y puede desempeñar un papel en la proliferación celular.
- *CIC*: la proteína codificada por este gen es un miembro de la superfamilia de represores transcripcionales del grupo de alta movilidad (HMG-box).
- *ARHGAP32*: codifica a una proteína activadora de GTPasa asociada a neuronas que puede regular la morfología y la fuerza de la columna dendrítica mediante la modulación de la actividad de Rho GTPasa.

5. DISCUSIÓN

5.1. ANÁLISIS DE LA METILACIÓN m⁶A

Numerosos estudios han demostrado que la desregulación de la modificación epigenética m⁶A está estrechamente relacionada con diversas enfermedades humanas, incluidos los tumores. Las alteraciones de los niveles de m⁶A pueden tener una profunda influencia en procesos celulares implicados en mantener la señalización proliferativa, evadir los efectos de los supresores del crecimiento, resistir la muerte celular, permitir la inmortalidad replicativa, inducir la angiogénesis, activar la invasión y la metástasis, reprogramar el metabolismo energético, la inestabilidad y mutación genómica, evadir la destrucción inmunitaria y promover la inflamación tumoral (58,202–206). Por lo tanto, la metilación de m⁶A puede desempeñar un papel importante en el cáncer. Por este motivo, en el presente estudio exploramos el estado de la metilación m⁶A, así como los niveles de expresión de los principales genes implicados en la metilación y desmetilación en el cáncer de mama en general y en los subtipos moleculares luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo.

Nuestros resultados mostraron que el porcentaje de metilación m⁶A global es significativamente menor en las muestras de tejido tumoral en comparación con sus correspondientes tejidos sanos. Además, el análisis de la expresión de los principales genes modificadores de m⁶A analizados, es decir, *METTL3*, *METTL14*, *WTAP*, *FTO* y *ALKBH5*, mostró que estos genes presentan una expresión reducida en muestras tumorales con respecto a sus correspondientes tejidos sanos. Por tanto, cabe suponer que el bajo porcentaje de metilación m⁶A global encontrado en las muestras tumorales podría deberse principalmente a la falta de metilación por parte del complejo de metiltransferasas. Este bajo porcentaje de metilación m⁶A global parece no poder ser compensado por la expresión también reducida de los genes que codifican las desmetilasas.

Los resultados obtenidos en estudios previos sobre la expresión génica de las metiltransferasas y desmetilasas informaron diferencias en la expresión (Tabla 1.3.1). Wu *et al.* (2019) analizaron la expresión génica utilizando los datos de las bases de datos ONCOMINE y TCGA y muestras de tejido mamario de 36 pacientes con cáncer de mama y encontraron una reducción significativa de la expresión en *METTL3*, *METTL14*, *WTAP* y *FTO* mientras que la expresión de *ALKBH5* se encontraba incrementada (125). En este caso, la expresión elevada de *ALKBH5* explicaba el bajo porcentaje de metilación m⁶A global encontrado en sus muestras tumorales. Por otro lado, Niu *et al.* (2019) informaron que el bajo porcentaje de metilación m⁶A global en sus muestras se debía al aumento de la expresión de *FTO* (127). Lv *et al.* (2021) observaron

que los niveles de expresión de *METTL14*, *WTAP*, *FTO* y *ALKBH5* estaban disminuidos en las muestras tumorales (129), mientras que Yang *et al.* (2.021) (128) y Qin *et al.* (2.021) (130) encontraron de forma independiente que la expresión de *METTL14*, *WTAP* y *FTO* se encontraba reducida, al igual que ocurre en nuestra muestra de pacientes. El estudio realizado por Yi *et al.* (2.020) (126) mostró que la expresión de *METTL14* se encontraba incrementada en las muestras tumorales de 20 pacientes con cáncer de mama. Por el contrario, Gong *et al.* (2.020) (131) observaron un bajo nivel de expresión de *METTL14* y *METTL3*, mientras que en el caso de Zheng *et al.* (2.021) (132), la expresión reducida se encontraba en *METTL14* y *WTAP*. En resumen, consideramos que las posibles razones de estas discrepancias en la expresión génica de las metiltransferasas y desmetilasas se deben a diferencias en el número y tipo de muestras analizadas en los estudios, al método utilizado para determinar la expresión génica y al hecho de que las muestras fueron analizadas sin tener en cuenta su subtipo molecular (132).

El cáncer de mama es una enfermedad heterogénea, que se clasifica en distintos subtipos moleculares (luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo) según el perfil de expresión génica (21,207). Por tanto, analizamos nuestros resultados de metilación m⁶A global teniendo en cuenta los diferentes subtipos moleculares. Encontramos que todos los subtipos coinciden en una menor metilación m⁶A, aunque solo presentó diferencias significativas entre las muestras tumorales y sus correspondientes tejidos sanos en el subtipo luminal A. Cuando analizamos la expresión de los genes que modifican la metilación m⁶A en los cinco subtipos, estos variaron entre sí (Figura 4.2.6-4.2.7, Tabla 4.2.4). Sin embargo, en todos los casos significativos, la expresión se encontraba reducida en las muestras tumorales en comparación con la expresión de los tejidos sanos. La expresión de *FTO* fue significativamente más baja en las muestras tumorales que en las muestras de tejido sano en todos los subtipos, por lo que, se podría esperar que los niveles de m⁶A aumentaran. Sin embargo, nuestros resultados mostraron una reducción general de la metilación m⁶A en las muestras tumorales en comparación con el tejido sano. Una posible explicación podría ser el efecto limitado de *FTO* sobre la desmetilación del ARN. Mauer y Jaffrey (2.018) han argumentado que *FTO* no tiene actividad específica de sustrato para m⁶A porque su tasa de desmetilación hacia m⁶A es lenta (a diferencia de la tasa de desmetilación de m⁶Am) y su papel como desmetilasa m⁶A ha sido cuestionada (54). De hecho, en algunos estudios, no se detectaron aumentos en los niveles de metilación de m⁶A en el ARNm cuando se usaba un tejido *knockdown* o *knockout* de *FTO* (53,208). Estas consideraciones explicarían por qué las muestras con expresión génica de las

metiltransferasas reducidas tendrían niveles más bajos de m⁶A incluso si la expresión de *FTO* también se encuentra reducida.

La metilación de m⁶A en el ARNm se produce como resultado del complejo de metiltransferasa compuesto por las proteínas METTL3, METTL14 y WTAP (53). Aunque METTL3 es la única proteína del complejo con actividad metiltransferasa (45), las otras dos son necesarias para la reacción de metilación. En particular, WTAP funciona como una proteína adaptadora sin actividad catalítica y promueve la localización del heterodímero METTL3-METTL14 en motas nucleares (49). En ausencia de WTAP, la metilación del ARNm disminuye drásticamente porque esta proteína influye en la unión de la parte catalítica del complejo de metiltransferasa (METTL3) al ARN (49,208). Este efecto podría explicar la reducción en la metilación en muestras con baja expresión de *ALKBH5* y *FTO* (como ocurre en los subtipos luminal A y HER2+, en las que la expresión de WTAP es baja), incluso si existen estudios que muestran que la actividad reducida de *ALKBH5* podría conducir a niveles más altos de m⁶A (208).

Las muestras tumorales analizadas en el presente estudio presentaron niveles significativamente más bajos de m⁶A en comparación con sus correspondientes tejidos sanos. Esta disminución en la metilación de m⁶A se observa en presencia de una expresión disminuida de los genes que codifican tanto las metiltransferasas (*METTL3*, *METTL14* y *WTAP*) como las desmetilasas (*FTO* y *ALKBH5*). Esta misma observación se repite al analizar cada uno de los subtipos moleculares. En este sentido, los niveles de expresión de estos genes muestran variaciones entre los distintos subtipos moleculares, como también se observa en el análisis de correlación realizado para evaluar las interacciones entre ellos. Las correlaciones más frecuentes ocurren entre metiltransferasas y desmetilasas, excepto en el subtipo HER2+. De hecho, encontramos diferentes correlaciones estadísticamente significativas específicas para cada subtipo molecular (Figura 4.2.10).

En resumen, los resultados actuales indican que, por un lado, a pesar de estar correlacionados, la reducción en la expresión de los genes que desmetilan no anula la reducción en la expresión de las metiltransferasas. Como resultado, encontramos una desmetilación neta de m⁶A. Por otro lado, cada subtipo molecular presenta un perfil de expresión génica de metiltransferasa y desmetilasas característico, así como un perfil de correlaciones, lo que podría ser útil para la clasificación molecular de tumores. Estos resultados destacan la necesidad de analizar cada subtipo en profundidad para entender mejor la modulación de la metilación m⁶A y su

implicación en el desarrollo de tumorigénesis en los diferentes subtipos moleculares de cáncer de mama.

5.2. IDENTIFICACIÓN *IN SILICO* Y ANÁLISIS DE EXPRESIÓN *IN VITRO* DE m⁶A-SNPs RELACIONADOS CON EL CÁNCER DE MAMA

Recientemente, los investigadores están comenzando a considerar los m⁶A-SNPs como una clase importante de variantes genéticas funcionales que pueden proporcionar nuevos conocimientos sobre la aparición, el desarrollo y la progresión de las enfermedades. En este sentido, los m⁶A-SNPs pueden estar involucrados en las enfermedades a través de múltiples efectos, por ejemplo, pueden afectar los niveles de metilación m⁶A, afectar la unión de factores de transcripción y proteínas, modificar la estructura secundaria del ARNm, y afectar la producción de ARNm maduro con función biológica correcta cuando se encuentran en los sitios de empalme (147,155). Por este motivo, se han realizado varios estudios con el fin de identificar los m⁶A-SNPs y evaluar el papel que desempeñan en diferentes enfermedades (147,148,159,160,150–152,154–158).

En este estudio, encontramos 981 m⁶A-SNPs después de comparar la lista de SNPs relacionados con el cáncer de mama con la lista de m⁶A-SNPs. De estos, solo cuatro m⁶A-SNPs presentaron un efecto *cis*-eQTL: rs1801270 en *CDKN1A*, rs76563149 en *ZNF354A*, rs9379084 en *REEB1* y rs11614913 en *MIR196A2*, y todos ellos, excepto rs9379084, mostraron una expresión génica alterada en el tejido mamario al utilizar los datos de las bases de datos GEO, ENCORI y de Wei *et al.* (2.015) (196). Estos m⁶A-SNPs son diferentes a los encontrados por Xuan *et al.* (2.021) (160), por lo que nuestro estudio permite ampliar la lista de m⁶A-SNPs que podrían estar involucrados en el cáncer de mama.

Para comprender el papel potencial de los m⁶A-SNPs que presentaron un efecto *cis*-eQTL y sus correspondientes genes en el cáncer de mama, investigamos cambios en la expresión de estos genes de manera *in silico*. El m⁶A-SNP rs9379084 se encuentra presente en el gen *RREB1*. Este gen no mostró diferencias en la expresión génica tras analizar los datos de la base de datos GEO. La proteína codificada por este gen actúa en diferentes procesos biológicos como la reparación del ADN dañado, el crecimiento y proliferación celular, la diferenciación celular, el desarrollo de grasas, equilibrio de glucosa en ayunas, transporte de zinc y la regulación transcripcional (209), por lo que podría desempeñar un papel en el cáncer. De hecho, *RREB1* se ha visto

involucrado en el desarrollo de varios cánceres y otras enfermedades (209), lo que indicaría que también podría estar relacionado con el desarrollo del cáncer de mama, aunque los efectos de la metilación de este SNP estarían implicados en procesos posteriores a la transcripción del gen.

El m⁶A-SNP rs11614913 se encuentra en el gen *MIR196A2*. Este gen codifica dos miARN maduros diferentes, miR-196a-5P y miR-196a-3P, del mismo tallo-bucle y es en el miR-196a-3P donde se encuentra el rs11614913. Este miARN maduro regula la expresión de múltiples genes diana, principalmente al emparejarse con la región 3'-UTR del ARNm, y afecta la estabilidad, inhibición de la traducción y/o la degradación del ARNm. De hecho, en diferentes condiciones, este miARN puede funcionar como un oncomiR (210). Este m⁶A-SNP ha sido genotipado en las pacientes con cáncer de mama y no se observaron diferencias significativas al compararlo con la población IBS (Tabla 4.3.3). Estos resultados concuerdan con los obtenidos por otros autores en poblaciones europeas (197,211,212). Sin embargo, el análisis de las poblaciones asiáticas muestra asociaciones significativas donde la variante C aumenta el riesgo de cáncer de mama (197–201,210). Por tanto, creemos que es importante evaluar el papel de este m⁶A-SNP en diferentes grupos étnicos. Por otro lado, como se muestra en la Figura 4.3.2, la presencia de la variante C o U no altera la estructura secundaria de este miARN, y dado que este m⁶A-SNP está cerca de cuatro posiciones m⁶A, el trabajo futuro debería abordar cómo este m⁶A-SNP puede afectar la modificación m⁶A y su influencia en el cáncer de mama.

El m⁶A-SNP rs1801270 está presente en el gen *CDKN1A*. La expresión de *CDKN1A* está controlada por la proteína supresora de tumores p53 y, por tanto, se relaciona con la detención del ciclo celular cuando el ADN se encuentra dañado. La proteína codificada por este gen está involucrada en múltiples funciones. Por un lado, la proteína CDKN1A se une e inhibe la actividad de las quinasas dependientes de ciclina, funcionando como un regulador de la progresión del ciclo celular en G1 cuando el ADN está dañado (213,214). Por otro lado, esta proteína interactúa con el antígeno nuclear de células en proliferación (PCNA) y desempeña un papel regulador en la replicación del ADN en fase S y la reparación del ADN dañado. También juega un papel en la apoptosis (215). Dada la participación de CDKN1A en estos procesos celulares, varias investigaciones han estudiado su papel en el desarrollo, progresión y terapia del cáncer (216,217). En este sentido, en comparación con sus correspondientes tejidos adyacentes no cancerosos, las muestras de cáncer de mama muestran un aumento de la expresión de CDKN1A (196). Por otro lado, varios estudios han asociado el rs1801270 con diferentes tipos de cáncer

y han informado que la presencia del alelo A es un factor de riesgo (218–223). Sin embargo, los estudios realizados en cáncer de mama en poblaciones europeas (224,225) no confirmaron su asociación con la variante A, cuya frecuencia en esta población es la más baja (Tabla 4.3.2). Esto se corresponde con los resultados experimentales obtenidos en el presente estudio. En nuestras pacientes, la presencia del alelo A no ha mostrado diferencias significativas con la población IBS. Por lo tanto, este m⁶A-SNP debe evaluarse en otras poblaciones donde su frecuencia es mayor, por ejemplo, en las poblaciones EAS, AFR o AMR, con el fin de determinar su participación en el cáncer de mama. A pesar de no encontrar diferencias, planteamos la hipótesis de que, dado que el m⁶A-SNP rs1801270 es una variante sin sentido que se encuentra en la región codificante, este m⁶A-SNP no solo podría tener un cambio de aminoácido, sino que podría afectar el procesamiento del ARN al alterar la estabilidad, la exportación nuclear, elongación de la transcripción o empalme del ARNm (72,144,147,148). Además, este m⁶A-SNP podría regular la traducción reduciendo la cinética del emparejamiento codón-anticodón (72) o resolviendo las estructuras secundarias del ARNm (63,72).

El m⁶A-SNP rs76563149, ubicado en el gen *ZNF354A*, pertenece a las proteínas de dominio de dedos de zinc asociadas a la caja Krüppel (KRAB-ZNF), una de las familias más abundantes de represores epigenéticos (226). Las KRAB-ZFP silencian la transcripción desencadenando la formación de heterocromatina y metilando las secuencias de ADN adyacentes a sus motivos de unión (226). Hasta la fecha, el papel de rs76563149 no ha sido explorado. Este m⁶A-SNP forma parte de un motivo DRACH y, en consecuencia, su presencia podría influir en la metilación de m⁶A. Genotipamos este m⁶A-SNP en las muestras de pacientes con cáncer de mama y nuestros resultados han mostrado una mayor frecuencia del alelo A en las pacientes que en las mujeres de la población IBS (Tabla 4.3.3). Por tal motivo, llevamos a cabo un análisis de expresión génica y encontramos bajos niveles de expresión en los tejidos mamarios tumorales en comparación con sus correspondientes tejidos mamarios sanos. Cuando realizamos el análisis de expresión génica teniendo en cuenta los subtipos moleculares, el subtipo luminal B HER2+ fue el único que presentó diferencias significativas. Estos resultados sugieren que el alelo A de este SNP podría alterar la capacidad de metilación m⁶A y afectar la expresión de *ZNF354A*. En este sentido, podría afectar el empalme, la exportación nuclear y la estabilidad del ARNm, así como la unión de las proteínas lectoras (59). De hecho, se ha observado que este m⁶A-SNP podría alterar el sitio de unión de los factores de transcripción y afectar a la unión de 11 proteínas (Tabla 4.3.1). Además, dado que este m⁶A-SNP se encuentra en la región 5'-UTR, puede

desempeñar un papel importante en la promoción de su traducción mediante un mecanismo de traducción independiente de la caperuza (36).

Aunque en este estudio nos hemos centrado en los m⁶A-SNPs que pueden influir en la expresión génica y son estadísticamente significativos, no hay que descartar el resto de m⁶A-SNPs que no presentaron estas características. La metilación m⁶A podría tener efectos más sutiles que implican una fina modulación de la expresión génica, por ejemplo, en el proceso de traducción y, en este caso, la influencia de la metilación no podría observarse mediante el análisis de eQTL. Por esta razón, se deben realizar estudios experimentales para confirmar si los 978 m⁶A-SNPs restantes podrían estar involucrados en el cáncer de mama.

En resumen, se identificaron tres m⁶A-SNPs tras el análisis *in silico* que podrían estar asociados con el cáncer de mama. Estos m⁶A-SNP, rs11614913, rs1801270 y rs76563149 se encuentran en los genes *MIR196A2*, *CDKN1A* y *ZNF354A*, respectivamente. Después del análisis *in vitro*, solo el m⁶A-SNP rs76563149 mostró diferencias significativas. Las futuras investigaciones en estos m⁶A-SNP deberían ampliar el estudio en diferentes grupos étnicos y aumentar el tamaño muestral para probar su asociación con el cáncer de mama y dilucidar su función molecular.

5.3. DETERMINACIÓN DE POSICIONES m⁶A METILADAS

Numerosas investigaciones realizadas durante los últimos años han demostrado que los distintos subtipos moleculares del cáncer de mama presentan diferentes patrones de metilación en el ADN (113,114,116,117). Sin embargo, hasta la fecha, el papel de la modificación m⁶A en el ARN de los diferentes subtipos moleculares de cáncer de mama aún no ha sido explorado.

En el presente estudio, hemos determinado el perfil de metilación m⁶A presente en todo el transcriptoma de muestras tumorales pertenecientes a los cinco subtipos moleculares de cáncer de mama mediante el uso de la técnica MeRIP-seq. El uso de ARN procedente de muestras de tejido tumoral supone un gran avance, ya que, estudios previos en el cáncer de mama mediante MeRIP-seq se habían realizado en cultivos celulares. Además, hemos realizado un análisis de GO y de las rutas metabólicas de KEGG para revelar las funciones potenciales de los primeros 20 genes que presentaban el *peak score* más elevado en cada subtipo molecular y hemos evaluado la expresión génica *in silico* de los primeros cinco genes. Hasta donde sabemos,

este es el primer estudio que describe el perfil de metilación m⁶A presente en los distintos subtipos moleculares.

La metodología MeRIP-seq fue la primera técnica desarrollada que permitió identificar las posiciones m⁶A (33,34). Posteriormente, otras técnicas fueron desarrollándose con el fin de solucionar las limitaciones que presentaba MeRIP-seq en cuando a la resolución y a la identificación de las posiciones m⁶A metiladas cuando hay dos o más motivos DRACH en el mismo fragmento inmunoprecipitado. En este sentido, cuando iniciamos este estudio (año 2.017), nuestro objetivo era identificar las posiciones m⁶A presente en los subtipos moleculares utilizando la metodología miCLIP dado que presentaba una resolución de un nucleótido. Sin embargo, tuvimos varios inconvenientes a la hora de la puesta a punto y optimización. Por un lado, nos encontramos con una falta de descripción detallada de los protocolos. Por el otro, la metodología miCLIP requiere equipamientos específicos no disponibles en la mayoría de los laboratorios. En este caso, para evitar trabajar con ³²P adoptamos la modificación propuesta por Zarnegar *et al.* (2.016) (110) sustituyendo el ³²P por un adaptador conjugado con un colorante infrarrojo. Sin embargo, el campus universitario de Álava carece del equipamiento para detectar el marcador infrarrojo. Por lo que, esto nos limitaba a la hora de llevar a cabo la puesta a punto y nos imposibilitaba continuar de inmediato con la siguiente etapa del protocolo. Además, cabe mencionar que la metodología miCLIP se caracteriza por estar compuesta por diversas técnicas tradicionales y habitualmente utilizadas en biología molecular, por lo que previamente requería una fase de puesta a punto individual de dichas técnicas. Por tal motivo, consideramos que las ventajas de la técnica meRIP-seq, cuyo protocolo es más sencillo y más corto y que utiliza materiales y equipamiento asequibles en la mayoría de los laboratorios, son superiores a sus limitaciones.

Los resultados obtenidos mediante la técnica MeRIP-seq revelaron un total de 1.703 picos m⁶A metilados. Estos variaban en función del subtipo, siendo el subtipo luminal B HER2- el que presentó más picos con 873 y el subtipo HER2+ el que menos con 55. Por otra parte, entre un 61,20-85,10% de los picos m⁶A identificados fueron específicos de cada subtipo molecular. También encontramos que, en todos los subtipos, la gran mayoría de los genes presentaban un único pico m⁶A, lo cual está en concordancia con los resultados de metilación m⁶A obtenidos mediante la técnica MeRIP-seq en otros tipos de cáncer (161-164).

Los picos m⁶A encontrados en los distintos subtipos tumorales se distribuyen principalmente en los exones y en la región 3'-UTR, por lo que podrían afectar a la expresión génica de

diferentes maneras. En el caso de los exones, la metilación m⁶A podría tener un efecto en el empalme alternativo (34) y podría afectar la traducción al reducir la cinética del emparejamiento codón-anticodón y resolviendo estructuras secundarias del ARNm mediante la unión de la proteína de unión al m⁶A YTHDC2 que contiene actividad ARN helicasa (63,72). En la región 3'-UTR, la metilación m⁶A podría afectar la poliadenilación alternativa e inhibir la transcripción mediada por miARN (33,38). Por otra parte, en el caso de subtipo HER2+, también se encontró un enriquecimiento de picos m⁶A en los intrones, los cuales se ha observado que podrían afectar el empalme alternativo en tres maneras distintas (227). Por un lado, la modificación m⁶A en intrones podría afectar la interacción entre el ARN pequeño nuclear (ARNpn) y el pre-ARNm. Por el otro, la modificación m⁶A en los intrones podría regular directamente la unión de proteínas de unión al ARN, y, por último, la modificación m⁶A podría afectar indirectamente los sitios de empalme al alterar la estructura secundaria del ARN.

También se observó que, en todos los subtipos moleculares, la mayoría de los picos m⁶A se encontraban en los cromosomas 1 y 19. En este sentido, la base de datos Cancer Genetics Web (<http://www.cancer-genetics.org/>) muestra que el cromosoma 1 posee el mayor número de genes relacionados con el cáncer de mama, con un total de 114 genes. Por lo que el m⁶A podría estar presente en los transcritos de dichos genes y, en consecuencia, el cromosoma 1 presenta un elevado N^o de picos m⁶A. En el caso del cromosoma 19, se encuentra en el puesto N^o 8 con 67 genes relacionados al cáncer de mama. Por otra parte, en el subtipo HER2+ no se ha observado picos m⁶A en los cromosomas 8,13 y X por lo que, futuros estudios debieran de evaluar como la falta de metilación m⁶A en estos cromosomas podría afectar a este subtipo molecular.

El análisis de GO y de las rutas metabólicas de KEGG se llevaron a cabo para identificar las funciones biológicas enriquecidas de los 20 primeros genes que presentaron el *peak score* más elevado de cada subtipo molecular. Nuestros resultados revelaron que los genes estaban relacionados con una variedad de procesos biológicos, componentes celulares y funciones moleculares distintas para cada subtipo. El análisis de rutas metabólicas mostró que únicamente los genes presentes en el subtipo HER2+ se encontraban asociados con la vía de señalización de la insulina. En este sentido, se ha observado que la insulina influye en el cáncer de mama de múltiples maneras. Por una parte, la insulina se une al receptor de insulina (RI) y activa la señalización PI3K/AKT/mTOR, promoviendo la proliferación, resistencia a la apoptosis y la invasión de las células tumorales (228). Por otra parte, los cánceres de mama

agresivos (como es el caso del subtipo HRE2+) se caracterizan por un alto consumo de glucosa y porque generan una mayor proporción de su energía a través de la glucólisis aeróbica (efecto Warburg). Esto se produce por la desregulación de la señalización PI3K/AKT/mTOR, la cual es activada por la insulina (228). También, se ha observado que la expresión de los RI se encuentra aumentada en los tejidos de cáncer de mama. A estos receptores, además de unirse a la insulina, se puede unir al factor de crecimiento 1 similar a la insulina (IGF1), que estimula las vías de señalización mitogénica (229). Por lo que, en un futuro, sería de gran interés estudiar todos los genes que presentaron picos m⁶A en este subtipo molecular y que se encontraron asociados a la vía de señalización de la insulina con el fin de identificar posibles dianas terapéuticas.

En cuanto al análisis de expresión génica *in silico* de los primeros cinco genes de la lista top 20 de cada subtipo, observamos que la gran mayoría presenta diferencias significativas en relación al tejido normal. Por lo que, la metilación m⁶A podría tener un efecto potencial en la transcripción. El gen *EPPK1* fue el único que estaba presente en la mayoría de los subtipos moleculares, a excepción del subtipo HER2+, y en todos los casos, se encontraba diferencialmente expresado. La proteína codificada por este gen pertenece a la familia de proteínas plakin que desempeñan un papel en la organización de la arquitectura citoesquelética. Diferentes estudios se han llevado a cabo sobre *EPPK1* en el carcinoma hepatocelular (230), el carcinoma urotelial (231), carcinoma cervical (232,233) y cáncer de páncreas (234). Sin embargo, el papel que desempeña *EPPK1* en el cáncer de mama aún permanece desconocido. Con respecto al resto de genes que presentaron diferencias en la expresión con el tejido normal, solo algunos de ellos se han estudiado en el cáncer de mama: *CBX4* (235), *IGFBP5* (236–239), *PHLDA1* (240–242), *FOXO3* (243), *MAGI1* (244), *CIC* (245) y *ARHGAP32* (246). Sin embargo, estos estudios se han realizado sin tener en cuenta la clasificación molecular. Por lo que, es de gran interés el estudio de la expresión de estos genes teniendo en cuenta el subtipo molecular en el que mostraron picos m⁶A.

En resumen, este es el primer estudio que describe el perfil de metilación m⁶A presente en todo el transcriptoma de los distintos subtipos moleculares de cáncer de mama obtenidos mediante la técnica MeRIP-seq. Además, proporcionamos sus posibles funciones biológicas y rutas metabólicas de los genes en donde se encuentran y evaluamos el efecto que la metilación m⁶A podría tener en la expresión génica para cada subtipo molecular. Estos resultados pueden ayudar a investigar más a fondo los mecanismos de regulación de la expresión génica mediada

por m⁶A, como también, la identificación de desarrollo de nuevas dianas terapéuticas para cada subtipo molecular.

6. CONCLUSIONES

1. La cuantificación de metilación m⁶A global de las muestras de cáncer de mama reveló que presentaban un bajo porcentaje de metilación en comparación con el tejido sano. Del mismo modo, se observó una baja expresión en los genes de las metiltransferasas y desmetilasas. Por tanto, cabe suponer que el bajo porcentaje de metilación m⁶A global encontrado en las muestras tumorales podría deberse principalmente a la falta de metilación por parte del complejo de metiltransferasas. En este sentido, la reducción en la expresión de los genes que desmetilan no anula la reducción en la expresión de las metiltransferasas y, como resultado, se observa una desmetilación neta de m⁶A.
2. Cada subtipo molecular presentó un perfil de expresión génica de metiltransferasa y desmetilasas, así como un perfil de correlaciones entre ellas característicos, lo que podría ser útil para la clasificación molecular de estos subtipos tumorales. Estos resultados destacan la necesidad de analizar cada subtipo en profundidad para entender mejor la modulación de la metilación m⁶A y su implicación en el desarrollo de la tumorigénesis en los diferentes subtipos moleculares de cáncer de mama.
3. Se han identificado tres m⁶A-SNPs que podrían estar asociados con el cáncer de mama tras el análisis *in silico*. Estos m⁶A-SNPs fueron: rs11614913, rs1801270 y rs76563149, situados en los genes *MIR196A2*, *CDKN1A* y *ZNF354A*, respectivamente. El m⁶A-SNP rs1801270, muestra baja frecuencia del alelo A, posiblemente asociado con la metilación m⁶A, en poblaciones europeas. Asimismo, la variante C del m⁶A-SNP rs11614913 se encuentra asociada al cáncer de mama en población asiática. Por ello, el estudio *in vitro* de estos m⁶A-SNPs requiere el uso de muestras de cáncer de mama de origen no europeo.
4. Únicamente el m⁶A-SNP rs76563149 ha mostrado diferencias significativas al comparar las frecuencias alélicas de las pacientes con cáncer de mama con la población IBS. Asimismo, se observaron diferencias significativas cuando se evaluó la expresión del gen *ZNF354A* en el conjunto total de muestras de tejido tumoral y tejido sano de las pacientes, y en el subtipo luminal B HER2+ en particular. Futuras investigaciones en este m⁶A-SNP, con un mayor tamaño muestral, permitirán evaluar su interés como marcador epitranscriptómico.

5. Se ha empleado la técnica MeRIP-seq para determinar el perfil de metilación m⁶A presente en los distintos subtipos moleculares de cáncer de mama. Ello ha permitido poner de manifiesto 1.703 picos m⁶A metilados y determinar picos m⁶A, tanto específicos de cada subtipo molecular como compartidos, el número de picos por transcritos, la distribución de los picos m⁶A en todo el transcriptoma y en los cromosomas. Este es el primer estudio de este tipo llevado a cabo en los distintos subtipos moleculares de cáncer de mama, por lo que, estos resultados generan nuevas vías de estudios en esta enfermedad.
6. El análisis de rutas metabólicas ha mostrado que las posiciones m⁶A metiladas presentes en el top 20 del subtipo HER2+ se encontraban en genes asociados a la vía de señalización de la insulina. Por lo que, sería de gran interés el análisis de la expresión de estos genes en este subtipo molecular con el fin de identificar posibles dianas terapéuticas.
7. El análisis de expresión *in silico* ha revelado que la gran mayoría de los genes que mostraron picos m⁶A metilados presentan diferencias significativas entre las muestras tumorales y el tejido normal. En este sentido, la metilación m⁶A podría ser la responsable de estas diferencias en la expresión. Por lo que, resulta de gran interés el estudio *in vitro* de la expresión de estos genes teniendo en cuenta el subtipo molecular en el que mostraron picos m⁶A.

7. BIBLIOGRAFÍA

1. Pró EA. Anatomía clínica. 1st ed. Argentina: Panamericana Medica, Editorial; 2012. 438–443 p.
2. Pawlina W. Ross. Histología: Texto y Atlas. 7a. Wolters Kluwer Health; 2015. 929–935 p.
3. Drake RL, Vogl W, Mitchell AWM. Gray anatomia para estudiantes. España: Elsevier España; 2005. 115–116 p.
4. Hassiotou F, Geddes D. Anatomy of the human mammary gland: Current status of knowledge. *Clin Anat.* 2013;26(1):29–48.
5. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394–424.
6. Momenimovahed Z, Salehiniya H. Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer (Dove Med Press).* 2019;11:151–64.
7. Zelli V, Compagnoni C, Capelli R, Cannita K, Sidoni T, Ficorella C, et al. Circulating micrnas as prognostic and therapeutic biomarkers in breast cancer molecular subtypes. *J Pers Med.* 2020;10(3):98.
8. Mueller C, Haymond A, Davis JB, Williams A, Espina V. Protein biomarkers for subtyping breast cancer and implications for future research. *Expert Rev Proteomics.* 2018;15(2):131–52.
9. Godone RLN, Leitão GM, Araújo NB, Castelletti CHM, Lima-Filho JL, Martins DBG. Clinical and molecular aspects of breast cancer: Targets and therapies. *Biomed Pharmacother.* 2018;106:14–34.
10. Tsang JYS, Tse GM. Molecular classification of breast cancer. *Adv Anat Pathol.* 2020;27(1):27-35.
11. Cserni G. Histological type and typing of breast carcinomas and the WHO classification changes over time. *Pathologica.* 2020;112(1):25–41.
12. Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, et al. Breast cancer prognostic classification in the molecular era: The role of histological grade. *Breast Cancer Res.* 2010;12(4):207.

13. Cserni G, Chmielik E, Cserni B, Tot T. The new TNM-based staging of breast cancer. *Virchows Arch.* 2018;472(5):697–703.
14. Breastcancer.org. Invasión de los sistemas vascular o linfático [Internet]. Breascancer.org. 2013. Available from: https://www.breastcancer.org/es/sintomas/diagnostico/invasion_vascular_linfatico
15. Rojas Duarte M, Marinkovic Gómez B. Cirugía en medicina general. 1o. Santiago, Chile: Centro de Enseñanza y Aprendizaje, Facultad de Medicina, Universidad de Chile; 2020. 407 p.
16. Jürgensen C M, Chacón C R, Baeza R C, Riveros P R. Cánceres de mama multifocales-multicéntricos: ¿Son realmente de peor pronóstico? *Rev Chil Cir.* 2009;61(2):125–30.
17. Kanumuri P, Hayse B, Killelea BK, Chagpar AB, Horowitz NR, Lannin DR. Characteristics of multifocal and multicentric breast cancers. *Ann Surg Oncol.* 2015;22(8):2475–82.
18. Tata A, Woolman M, Ventura M, Bernards N, Ganguly M, Gribble A, et al. Rapid detection of necrosis in breast cancer with desorption electrospray ionization mass spectrometry. *Sci Rep.* 2016;6:35374.
19. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. *Nat Rev Dis Prim.* 2019;5(1):66.
20. Perou CM, Sùrlie T, Eisen MB, Rijn M Van De, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature.* 2000;406:747–52.
21. Makki J. Diversity of breast carcinoma: Histological subtypes and clinical relevance. *Clin Med Insights Pathol.* 2015;8:23–31.
22. Horvath E. Subtipos moleculares del cáncer mamario - lo que el radiólogo dedicado a imágenes mamarias debe saber. *Rev Chil Radiol.* 2021;27(1):17–26.
23. Boccaletto P, MacHnicka MA, Purta E, Pitkowski P, Baginski B, Wirecki TK, et al. MODOMICS: A database of RNA modification pathways. *Nucleic Acids Res.* 2018;46(D1):D303–7.
24. Lewis CJT, Pan T, Kalsotra A. RNA modifications and structures cooperate to guide RNA-protein interactions. *Nat Rev Mol Cell Biol.* 2017;18(3):202–10.

25. Perry RP, Kelley DE. Existence of methylated messenger RNA in mouse L cells. *Cell*. 1974;1(1):37–42.
26. Desrosiers R, Friderici K, Rottman F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc Natl Acad Sci USA*. 1974;71(10):3971–5.
27. Lavi S, Shatkin AJ. Methylated simian virus 40 specific RNA from nuclei and cytoplasm of infected BSC 1 cells. *Proc Natl Acad Sci USA*. 1975;72(6):2012–6.
28. Wei CM, Moss B. Methylated nucleotides block 5' terminus of vaccinia virus messenger RNA. *Proc Natl Acad Sci USA*. 1975;72(1):318–22.
29. Furuichi Y, Morgan M, Shatkin AJ, Jelinek W, Salditt-Georgieff M, Darnell JE. Methylated, blocked 5' termini in HeLa cell mRNA. *Proc Natl Acad Sci USA*. 1975;72(5):1904–8.
30. Adams JM, Cory S. Modified nucleosides and bizarre 5'-termini in mouse myeloma mRNA. *Nature*. 1975;255(5503):28–33.
31. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, et al. N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol*. 2011;7(12):885–7.
32. Cao G, Li HB, Yin Z, Flavell RA. Recent advances in dynamic m6A RNA modification. *Open Biol*. 2016;6(4):160003.
33. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*. 2012;149(7):1635–46.
34. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*. 2012;485(7397):201–6.
35. Linder B, Grozhik A V, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods*. 2015;12(8):767–72.
36. Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, et al. 5' UTR m(6)A promotes cap-independent translation. *Cell*. 2015;163(4):999–1010.

37. Meyer KD, Jaffrey SR. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nat Rev Mol Cell Biol.* 2014;15(5):313–26.
38. Ke S, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, et al. A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* 2015;29(19):2037–53.
39. Berulava T, Rahmann S, Rademacher K, Klein-Hitpass L, Horsthemke B. N6-Adenosine methylation in miRNAs. *PLoS One.* 2015;10(2):e0118438.
40. Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, et al. Extensive translation of circular RNAs driven by N6-methyladenosine. *Cell Res.* 2017;27(5):626–41.
41. Pinello N, Sun S, Wong JLL. Aberrant expression of enzymes regulating m6A mRNA methylation: Implication in cancer. *Cancer Biol Med.* 2018;15(4):323–34.
42. Yang Y, Hsu PJ, Chen YS, Yang YG. Dynamic transcriptomic m6A decoration: Writers, erasers, readers and functions in RNA metabolism. *Cell Res.* 2018;28(6):616–24.
43. Zhang W, Qian Y, Jia G. The detection and functions of RNA modification m6A based on m6A writers and erasers. *J Biol Chem.* 2021;297(2):100973.
44. Ru W, Zhang X, Yue B, Qi A, Shen X, Huang Y, et al. Insight into m6A methylation from occurrence to functions. *Open Biol.* 2020;10(9):200091.
45. Bokar JA, Shambaugh ME, Polayes D, Matera AG, Rottman FM. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA.* 1997;3(11):1233–47.
46. Wang P, Doxtader KA, Nam Y. Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Mol Cell.* 2016;63(2):306–17.
47. Śledź P, Jinek M. Structural insights into the molecular mechanism of the m6A writer complex. *Elife.* 2016;5:e18434.
48. Wang X, Feng J, Xue Y, Guan Z, Zhang D, Liu Z, et al. Structural basis of N6-adenosine methylation by the METTL3-METTL14 complex. *Nature.* 2016;534:575–8.

49. Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, et al. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 2014;24(2):177–89.
50. Liu J, Yue Y, Liu J, Cui X, Cao J, Luo G, et al. VIRMA mediates preferential m6A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov.* 2018;4:10.
51. Patil DP, Chen C-K, Pickering BF, Chow A, Jackson C, Guttman M, et al. m6A RNA methylation promotes XIST-mediated transcriptional repression. *Nature.* 2016;537(7620):369–73.
52. Bawankar P, Lence T, Paolantoni C, Hausmann IU, Kazlauskienė M, Jacob D, et al. Hakai is required for stabilization of core components of the m6A mRNA methylation machinery. *Nat Commun.* 2021;12(1):3778.
53. Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik A V., Patil DP, et al. Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature.* 2017;541(7237):371–5.
54. Mauer J, Jaffrey SR. FTO, m6Am, and the hypothesis of reversible epitranscriptomic mRNA modifications. *FEBS Lett.* 2018;592(12):2012–22.
55. Wei J, Fange L, Lu Z, Qili F, Yuxi A, He PC, et al. Differential m6A, m6Am, and m1A demethylation mediated by FTO in cell nucleus and cytoplasm. *Mol Cell.* 2018;71(6):973–85.
56. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Li CJ, Vågbø CB, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. 2013;49(1):18–29.
57. Toh JDW, Crossley SWM, Bruemmer KJ, Ge EJ, He D, Iovan DA, et al. Distinct RNA N-demethylation pathways catalyzed by nonheme iron ALKBH5 and FTO enzymes enable regulation of formaldehyde release rates. *Proc Natl Acad Sci USA.* 2020;117(41):25284–92.
58. Zhou Z, Lv J, Yu H, Han J, Yang X, Feng D, et al. Mechanism of RNA modification N6-methyladenosine in human cancer. *Mol Cancer.* 2020;19(1):104.

59. Berlivet S, Scutenaire J, Deragon JM, Bousquet-Antonelli C. Readers of the m6A epitranscriptomic code. *Biochim Biophys Acta - Gene Regul Mech.* 2019;1862(3):329–42.
60. Reichel M, Köster T, Staiger D. Marking RNA: m6A writers, readers, and functions in Arabidopsis. *J Mol Cell Biol.* 2019;11(10):899–910.
61. Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, et al. Nuclear m6A reader YTHDC1 regulates mRNA splicing. *Mol Cell.* 2016;61(4):507–19.
62. Zhao Y, Shi Y, Shen H, Xie W. m6A-binding proteins: The emerging crucial performers in epigenetics. *J Hematol Oncol.* 2020;13:35.
63. Mao Y, Dong L, Liu X-M, Guo J, Ma H, Shen B, et al. m6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat Commun.* 2019;10(1):5332.
64. Alarcón CR, Goodarzi H, Lee H, Liu X, Tavazoie S, Tavazoie SF. HNRNPA2B1 is a mediator of m6A-dependent nuclear RNA processing events. *Cell.* 2015;162(6):1299–308.
65. Alarcón CR, Lee H, Goodarzi H, Halberg N, Tavazoie SF. N6-methyladenosine marks primary microRNAs for processing. *Nature.* 2015;519(7544):482–5.
66. Huang H, Weng H, Sun W, Qin X, Shi H, Wu H, et al. Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat Cell Biol.* 2018;20(3):285–95.
67. Zhang Y, Geng X, Li Q, Xu J, Tan Y, Xiao M, et al. m6A modification in RNA: Biogenesis, functions and roles in gliomas. *J Exp Clin Cancer Res.* 2020;39(1):192.
68. Liu N, Zhou KI, Parisien M, Dai Q, Diatchenko L, Pan T. N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.* 2017;45(10):6051–63.
69. Zhao X, Yang Y, Sun BF, Shi Y, Yang X, Xiao W, et al. FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* 2014;24(12):1403–19.
70. Bartosovic M, Molares HC, Gregorova P, Hrossova D, Kudla G, Vanacova S. N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3'-end processing. *Nucleic Acids Res.* 2017;45(19):11356–70.

71. Roundtree IA, Luo GZ, Zhang Z, Wang X, Zhou T, Cui Y, et al. YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *Elife*. 2017;6:e31311.
72. Slobodin B, Dikstein R. So close, no matter how far: Multiple paths connecting transcription to mRNA translation in eukaryotes. *EMBO Rep*. 2020;21(9):e50799.
73. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N6-methyladenosine modulates messenger RNA translation efficiency. *Cell*. 2015;161(6):1388–99.
74. Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*. 2014;505(7481):117–20.
75. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature*. 2015;518(7540):560–4.
76. Wu B, Su S, Patil DP, Liu H, Gan J, Jaffrey SR, et al. Molecular basis for the specific and multivalent recognitions of RNA substrates by human hnRNP A2/B1. *Nat Commun*. 2018;9(1):420.
77. Han X, Guo J, Fan Z. Interactions between m6A modification and miRNAs in malignant tumors. *Cell Death Dis*. 2021;12(6):598.
78. Batista PJ, Molinie B, Wang J, Qu K, Zhang J, Li L, et al. m6A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell*. 2014;15(6):707–19.
79. Geula S, Moshitch-Moshkovitz S, Dominissini D, Mansour AAF, Kol N, Salmon-Divon M, et al. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science*. 2015;347(6225):1002–6.
80. Wang Y, Li Y, Toth JI, Petroski MD, Zhang Z, Zhao JC. N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat Cell Biol*. 2014;16(2):191–8.
81. Fustin JM, Doi M, Yamaguchi Y, Hida H, Nishimura S, Yoshida M, et al. RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell*. 2013;155(4):793–806.
82. Zhou J, Wan J, Gao X, Zhang X, Qian S-B. Dynamic m6A mRNA methylation directs translational control of heat shock response. *Nature*. 2015;526(7574):591–4.

83. Schwartz S, Agarwala SD, Mumbach MR, Jovanovic M, Mertins P, Shishkin A, et al. High-resolution mapping reveals a conserved, widespread, dynamic meiotically regulated mRNA methylation program. *Cell*. 2013;155(6):1409–21.
84. McIntyre ABR, Gokhale NS, Cerchietti L, Jaffrey SR, Horner SM, Mason CE. Limits in the detection of m6A changes using MeRIP/m6A-seq. *Sci Rep*. 2020;10(1):6590.
85. Dominissini D, Moshitch-Moshkovitz S, Salmon-Divon M, Amariglio N, Rechavi G. Transcriptome-wide mapping of N6-methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing. *Nat Protoc*. 2013;8(1):176–89.
86. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of CHIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.
87. Meng J, Cui X, Rao MK, Chen Y, Huang Y. Exome-based analysis for RNA epigenome sequencing data. *Bioinformatics*. 2013;29(12):1565–7.
88. Cui X, Meng J, Zhang S, Chen Y, Huang Y. A novel algorithm for calling mRNA m6A peaks by modeling biological variances in MeRIP-seq data. *Bioinformatics*. 2016;32(12):i378–85.
89. Zeng Y, Wang S, Gao S, Soares F, Ahmed M, Guo H, et al. Refined RIP-seq protocol for epitranscriptome analysis with low input materials. *PLoS Biol*. 2018;16(9):e2006092.
90. Linder B, Grozhik A V, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods*. 2015;12(8):767–72.
91. Grozhik A V, Linder B, Olarerin-George AO, Jaffrey SR. Mapping m6A at individual-nucleotide resolution using crosslinking and immunoprecipitation (miCLIP). *Methods Mol Biol*. 2017;1562:55–78.
92. Hawley BR, Jaffrey SR. Transcriptome-wide mapping of m6A and m6Am at single-nucleotide resolution using miCLIP. *Curr Protoc Mol Biol*. 2019;126(1):e88.
93. Chen K, Lu Z, Wang X, Fu Y, Luo GZ, Liu N, et al. High-resolution N6-methyladenosine (m6A) map using photo-crosslinking-assisted m6A sequencing. *Angew Chemie - Int Ed*. 2015;54(5):1587–90.

94. Meyer K. DART-seq: An antibody-free method for global m6A detection. *Nat Methods*. 2019;16(12):1275–80.
95. Leger A, Amaral PP, Pandolfini L, Capitanchik C, Capraro F, Barbieri I, et al. RNA modifications detection by comparative Nanopore direct RNA sequencing. *bioRxiv*. 2019;12(7198):1–29.
96. Lorenz DA, Sathe S, Einstein JM, Yeo GW. Direct RNA sequencing enables m6A detection in endogenous transcript isoforms at base-specific resolution. *RNA*. 2020;26(1):19–28.
97. Liu H, Begik O, Lucas MC, Ramirez JM, Mason CE, Wiener D, et al. Accurate detection of m6A RNA modifications in native RNA sequences. *Nat Commun*. 2019;10(1):4079.
98. Zhang Z, Chen L-Q, Zhao Y-L, Yang C-G, Roundtree I, Zhang Z, et al. Single-base mapping of m6A by an antibody-independent method. *Sci Adv*. 2019;5(7):eaax0250.
99. Koh CWQ, Goh YT, Goh WSS. Atlas of quantitative single-base-resolution N6-methyladenine methylomes. *Nat Commun*. 2019;10(1).
100. Wang Y, Xiao Y, Dong S, Yu Q, Jia G. Antibody-free enzyme-assisted chemical approach for detection of N6-methyladenosine. *Nat Chem Biol*. 2020;16(8):896–903.
101. Shu X, Cao J, Cheng M, Xiang S, Gao M, Li T, et al. A metabolic labeling method detects m6A transcriptome-wide at single base resolution. *Nat Chem Biol*. 2020;16(8):887–95.
102. Shen L, Liang Z, Yu H. Dot blot analysis of N6-methyladenosine RNA modification levels. *Bio-Protocol*. 2017;7(1):e2095.
103. William Weihao Tang, Matthew Perry Young, Anila Mamidi, Jose Angel Regla-Nava, Kenneth Kim SS. N6-methyladenosine RNA modification regulates shoot stem cell fate in Arabidopsis. *Physiol Behav*. 2017;176(3):139–48.
104. Wang CY, Lin MH, Su HT. A method for measuring RNA N6-methyladenosine modifications in cells and tissues. *J Vis Exp*. 2016;(118):54672.
105. Mishima E, Jinno D, Akiyama Y, Itoh K, Nankumo S, Shima H, et al. Immuno-northern blotting: Detection of RNA modifications by using antibodies against modified nucleosides. *PLoS One*. 2015;10(11):1–17.

106. Molinie B, Wang J, Lim KS, Hillebrand R, Lu ZX, Van Wittenberghe N, et al. m6A-LAIC-seq reveals the census and complexity of the m6A epitranscriptome. *Nat Methods*. 2016;13(8):692–8.
107. Garcia-Campos MA, Edelheit S, Toth U, Safra M, Shachar R, Viukov S, et al. Deciphering the “m6A code” via antibody-independent quantitative profiling. *Cell*. 2019;178(3):731–47.
108. Pandey RR, Pillai RS. Counting the cuts: MAZTER-Seq quantifies m6A levels using a methylation-sensitive ribonuclease. *Cell*. 2019;178(3):515–7.
109. Li X, Xiong X, Yi C. Epitranscriptome sequencing technologies: Decoding RNA modifications. *Nat Methods*. 2016;14(1):23–31.
110. Zarnegar BJ, Flynn RA, Shen Y, Do BT, Chang HY, Khavari PA. IrCLIP platform for efficient characterization of protein-RNA interactions. *Nat Methods*. 2016;13(6):489–92.
111. Stefansson OA, Esteller M. Epigenetic modifications in breast cancer and their role in personalized medicine. *Am J Pathol*. 2013;183(4):1052–63.
112. Basse C, Arock M. The increasing roles of epigenetics in breast cancer: Implications for pathogenicity, biomarkers, prevention and treatment. *Int J Cancer*. 2015;137(12):2785–94.
113. Bediaga NG, Acha-sagredo A, Guerra I, Viguri A, Albaina C, Diaz IR, et al. DNA methylation epigenotypes in breast cancer molecular subtypes. *Breast Cancer Res*. 2010;12(5):R77.
114. Bediaga NG, Beristain E, Calvo B, Viguri MA, Gutierrez-Corres B, Rezola R, et al. Luminal B breast cancer subtype displays a dicotomic epigenetic pattern. *Springerplus*. 2016;5:623.
115. Jin W, Li QZ, Liu Y, Zuo YC. Effect of the key histone modifications on the expression of genes related to breast cancer. *Genomics*. 2020;112(1):853–8.
116. Bao X, Anastasov N, Wang Y, Rosemann M. A novel epigenetic signature for overall survival prediction in patients with breast cancer. *J Transl Med*. 2019;17:380.
117. Chen X, Zhang J, Dai X. DNA methylation profiles capturing breast cancer heterogeneity. *BMC Genomics*. 2019;20:823.

118. Batista PJ. The RNA modification N6-methyladenosine and its implications in human disease. *Genomics, Proteomics & Bioinformatic*. 2017;15(3):154–63.
119. He L, Li J, Wang X, Ying Y, Xie H, Yan H, et al. The dual role of N6-methyladenosine modification of RNAs is involved in human cancers. *J Cell Mol Med*. 2018;22(10):4630–9.
120. Wang Y, Zhang Y, Du Y, Zhou M, Hu Y, Zhang S. Emerging roles of N6-methyladenosine (m6A) modification in breast cancer. *Cell Biosci*. 2020;10(1):136.
121. Huang H, Weng H, Chen J. m6A modification in coding and non-coding RNAs: Roles and therapeutic implications in cancer. *Cancer Cell*. 2020;37(3):270–88.
122. Li H, Wu H, Wang Q, Ning S, Xu S, Pang D. Dual effects of N6-methyladenosine on cancer progression and immunotherapy. *Mol Ther - Nucleic Acids*. 2021;24:25–39.
123. Wei M, Bai JW, Niu L, Zhang YQ, Chen HY, Zhang GJ. The complex roles and therapeutic implications of m6A modifications in breast cancer. *Front Cell Dev Biol*. 2021;8:615071.
124. Hu Y, Wang S, Liu J, Huang Y, Gong C, Liu J, et al. New sights in cancer: Component and function of N6-methyladenosine modification. *Biomed Pharmacother*. 2020;122:109694.
125. Wu L, Wu D, Ning J, Liu W, Zhang D. Changes of N6-methyladenosine modulators promote breast cancer progression. *BMC Cancer*. 2019;19(1):326.
126. Yi D, Wang R, Shi X, Xu L, Yilihamu Y, Sang J. METTL14 promotes the migration and invasion of breast cancer cells by modulating N6-methyladenosine and hsa-miR-146a-5p expression. *Oncol Rep*. 2020;43(5):1375–86.
127. Niu Y, Lin Z, Wan A, Chen H, Liang H, Sun L, et al. RNA N6-methyladenosine demethylase FTO promotes breast tumor progression through inhibiting BNIP3. *Mol Cancer*. 2019;18(1):46.
128. Yang L, Wu S, Ma C, Song S, Jin F, Niu Y, et al. RNA m6A methylation regulators subclassify luminal subtype in breast cancer. *Front Oncol*. 2021;10:611191.
129. Lv W, Wang Y, Zhao C, Tan Y, Xiong M, Yi Y, et al. Identification and validation of m6A-related lncRNA signature as potential predictive biomarkers in breast cancer. *Front Oncol*. 2021;11:745719.

130. Qin Q, Fang DL, Zhou W, Meng Y, Wei J. Classification and immune invasion analysis of breast cancer based on m6A genes. *Ann Transl Med.* 2021;9(18):1418.
131. Gong PJ, Shao YC, Yang Y, Song WJ, He X, Zeng YF, et al. Analysis of N6-methyladenosine methyltransferase reveals METTL14 and ZC3H13 as tumor suppressor genes in breast cancer. *Front Oncol.* 2020;10:578963.
132. Zheng F, Du F, Qian H, Zhao J, Wang X, Yue J, et al. Expression and clinical prognostic value of m6A RNA methylation modification in breast cancer. *Biomark Res.* 2021;9(1):28.
133. Fang R, Ye L, Shi H. Understanding the roles of N6-methyladenosine writers, readers and erasers in breast cancer. *Neoplasia.* 2021;23(6):551–60.
134. Wang H, Xu B, Shi J. N6-methyladenosine METTL3 promotes the breast cancer progression via targeting Bcl-2. *Gene.* 2020;722:144076.
135. Cai X, Wang X, Cao C, Gao Y, Zhang S, Yang Z, et al. HBXIP-elevated methyltransferase METTL3 promotes the progression of breast cancer via inhibiting tumor suppressor let-7g. *Cancer Lett.* 2018;415:11–9.
136. Rong D, Dong Q, Qu H, Deng X, Gao F, Li Q, et al. m6A-induced LINC00958 promotes breast cancer tumorigenesis via the miR-378a-3p/YY1 axis. *Cell Death Discov.* 2021;7(1):27.
137. Shi Y, Zheng C, Jin Y, Bao B, Wang D, Hou K, et al. Reduced expression of METTL3 promotes metastasis of triple-negative breast cancer by m6A methylation-mediated COL3A1 up-regulation. *Front Oncol.* 2020;10:1126.
138. Yi D, Wang R, Shi X, Xu L, Yilihamu Y, Sang J. METTL14 promotes the migration and invasion of breast cancer cells by modulating N6-methyladenosine and hsa-miR-146a-5p expression. *Oncol Rep.* 2020;43(5):1375–86.
139. Sun T, Wu Z, Wang X, Wang Y, Hu X, Qin W, et al. LNC942 promoting METTL14-mediated m6A methylation in breast cancer cell proliferation and progression. *Oncogene.* 2020;39(31):5358–72.

140. Zhang C, Samanta D, Lu H, Bullen JW, Zhang H, Chen I, et al. Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m6A-demethylation of NANOG mRNA. *Proc Natl Acad Sci USA*. 2016;113(14):E2047-2056.
141. Zhang C, Zhi WI, Lu H, Samanta D, Chen I, Gabrielson E, et al. Hypoxia-inducible factors regulate pluripotency factor expression by ZNF217-and ALKBH5-mediated modulation of RNA methylation in breast cancer cells. *Oncotarget*. 2016;7(40):64527–42.
142. Chang G, Shi L, Ye Y, Shi H, Zeng L, Tiwary S, et al. YTHDF3 induces the translation of m6A-enriched gene transcripts to promote breast cancer brain metastasis. *Cancer Cell*. 2020;38(6):857-871.e7.
143. Zhu P, He F, Hou Y, Tu G, Li Q, Jin T, et al. A novel hypoxic long noncoding RNA KB-1980E6.3 maintains breast cancer stem cell stemness via interacting with IGF2BP1 to facilitate c-Myc mRNA stability. *Oncogene*. 2021;40(9):1609–27.
144. Jiang S, Xie Y, He Z, Zhang Y, Zhao Y, Chen L, et al. m6ASNP: A tool for annotating genetic variants by m6A function. *Gigascience*. 2018;7(5):giy035.
145. Zheng Y, Nie P, Peng D, He Z, Liu M, Xie Y, et al. M6AVar: A database of functional variants involved in m6A modification. *Nucleic Acids Res*. 2018;46(D1):D139-145.
146. Qiyuan L, Seo J-H, Stranger B, McKenna A, Pe'er I, LaFramboise T, et al. A novel eQTL-based analysis reveals the biology of breast cancer risk loci. *Cell*. 2013;152(3):633–41.
147. Lin W, Xu H, Wu Y, Wang J, Yuan Q. In silico genome-wide identification of m6A-associated SNPs as potential functional variants for periodontitis. *J Cell Physiol*. 2020;235(2):900–8.
148. Sun X, Dai Y, Tan G, Liu Y, Li N. Integration analysis of m6A-SNPs and eQTLs associated with sepsis reveals platelet degranulation and *Staphylococcus aureus* infection are mediated by m6A mRNA methylation. *Front Genet*. 2020;11:7.
149. Chen M, Lin W, Yi J, Zhao Z. Exploring the epigenetic regulatory role of m6A-associated SNPs in type 2 diabetes pathogenesis. *Pharmgenomics Pers Med*. 2021;14:1369–78.
150. Ruan X, Tian M, Kang N, Ma W, Zeng Y, Gaojian Zhuang³ WZ, et al. Genome-wide identification of m6A-associated functional SNPs as potential functional variants for thyroid cancer. *Am J Cancer Res*. 2021;11(11):5402–14.

151. Mo XB, Zhang YH, Lei SF. Genome-wide identification of N6-methyladenosine (m6A) SNPs associated with rheumatoid arthritis. *Front Genet.* 2018;9:299.
152. Mo XB, Zhang YH, Lei SF. Genome-wide identification of m6A-associated SNPs as potential functional variants for bone mineral density. *Osteoporos Int.* 2018;29(9):2029–39.
153. Hu Y, Zhao X. Role of m6A in osteoporosis, arthritis and osteosarcoma (Review). *Exp Ther Med.* 2021;22:926.
154. Mo XB, Lei SF, Zhang YH, Zhang H. Detection of m6A-associated SNPs as potential functional variants for coronary artery disease. *Epigenomics.* 2018;10(10):1279–87.
155. Lin W, Xu H, Yuan Q, Zhang S. Integrative genomic analysis predicts regulatory role of N6-methyladenosine-associated SNPs for adiposity. *Front Cell Dev Biol.* 2020;8:551.
156. Mo X, Lei S, Zhang Y, Zhang H. Genome-wide enrichment of m6A-associated single-nucleotide polymorphisms in the lipid loci. *Pharmacogenomics J.* 2019;19(4):347–57.
157. Qiu X, He H, Huang Y, Wang J, Xiao Y. Genome-wide identification of m6A-associated single-nucleotide polymorphisms in Parkinson's disease. *Neurosci Lett.* 2020;737:135315.
158. Sebastian-delaCruz M, Olazagoitia-Garmendia A, Gonzalez-Moro I, Santin I, Garcia-Etxebarria K, Castellanos-Rubio A. Implication of m6A mRNA methylation in susceptibility to inflammatory bowel disease. *Epigenomes.* 2020;4(3):16.
159. Zhao H, Jiang J, Wang M, Xuan Z. Genome-wide identification of m6A-associated single-nucleotide polymorphisms in colorectal cancer. *Pharmgenomics Pers Med.* 2021;14:887–92.
160. Xuan Z, Zhang Y, Jiang J, Zheng X, Hu X, Yang X, et al. Integrative genomic analysis of N6-methyladenosine-single nucleotide polymorphisms (m6A-SNPs) associated with breast cancer. *Bioengineered.* 2021;12(1):2389–97.
161. Li A, Gan Y, Cao C, Ma B, Zhang Q, Zhang Q, et al. Transcriptome-wide map of N6-methyladenosine methylome profiling in human bladder cancer. *Front Oncol.* 2021;11:717622.

162. Chen Y, Zhou C, Sun Y, He X, Xue D. m6A RNA modification modulates gene expression and cancer-related pathways in clear cell renal cell carcinoma. *Epigenomics*. 2019;12(2):87–99.
163. Han Z, Yang B, Wang Q, Hu Y, Wu Y, Tian Z. Comprehensive analysis of the transcriptome-wide m6A methylome in invasive malignant pleomorphic adenoma. *Cancer Cell Int*. 2021;21(1):142.
164. Niu X, Xu J, Liu J, Chen L, Qiao X, Zhong M. Landscape of N6-methyladenosine modification patterns in human ameloblastoma. *Front Oncol*. 2020;10:556497.
165. Yang H, Wu Y-F, Ding J, Liu W, Zhu D-S, Shen X-F, et al. Comprehensive analysis of N6-methyladenosine (m6A) methylation in neuromyelitis optica spectrum disorders. *Front Genet*. 2021;12:735454.
166. Wen K, Zhang Y, Li Y, Wang Q, Sun J. Comprehensive analysis of transcriptome-wide m6A methylome in the anterior capsule of the lens of high myopia patients. *Epigenetics*. 2021;16(9):955–68.
167. Berulava T, Buchholz E, Elerdashvili V, Pena T, Islam MR, Lbik D, et al. Changes in m6A RNA methylation contribute to heart failure progression by modulating translation. *Eur J Heart Fail*. 2020;22(1):54–66.
168. Wang Y, Zeng L, Liang C, Zan R, Ji W, Zhang Z, et al. Integrated analysis of transcriptome-wide m6A methylome of osteosarcoma stem cells enriched by chemotherapy. *Epigenomics*. 2019;11(15):1693–715.
169. Wang J, Gao F, Zhao X, Cai Y, Jin H. Integrated analysis of the transcriptome-wide m6A methylome in preeclampsia and healthy control placentas. *PeerJ*. 2020;8:e9880.
170. Wang YY, Lehuédé C, Laurent V, Dirat B, Dauvillier S, Bochet L, et al. Adipose tissue and breast epithelial cells: A dangerous dynamic duo in breast cancer. *Cancer Lett*. 2012;324(2):142–51.
171. Rio DC, Ares M, Hannon GJ, Nilsen TW. Nondenaturing agarose gel electrophoresis of RNA. *Cold Spring Harb Protoc*. 2010;5(6):1–3.
172. Marshall OJ. PerlPrimer: Cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics*. 2004;20(15):2471–2.

173. Wang X, Li Z, Kong B, Song C, Cong J, Hou J, et al. Reduced m6A mRNA methylation is correlated with the progression of human cervical cancer. *Oncotarget*. 2017;8(58):98918–30.
174. Valente V, Teixeira SA, Neder L, Okamoto OK, Oba-Shinjo SM, Marie SKN, et al. Selection of suitable housekeeping genes for expression analysis in glioblastoma using quantitative RT-PCR. *Ann eurosciences*. 2014;21(2):62–3.
175. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;(13):134.
176. King N. Measurement of cardiac gene expression by reverse transcription polymerase chain reaction (RT-PCR). *Methods Mol Biol*. 2007;366:109–19.
177. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods*. 2001;25(4):402–8.
178. Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kibbe WA, et al. Toward a shared vision for cancer genomic data. *N Engl J Med*. 2016;375(12):1109–12.
179. Buniello A, Macarthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*. 2019;47(D1):D1005–12.
180. Cariaso M, Lennon G. SNPedia: A wiki supporting personal genome annotation, interpretation and analysis. *Nucleic Acids Res*. 2012;40(Database issue):D1308-1312.
181. Bruno AE, Li L, Kalabus JL, Pan Y, Yu A, Hu Z. miRdSNP: A database of disease-associated SNPs and microRNA target sites on 3'UTRs of human genes. *BMC Genomics*. 2012;13:44.
182. Liu C-J, Fu X, Xia M, Zhang Q, Gu Z, Guo A-Y. miRNASNP-v3: A comprehensive database for SNPs and disease-related variations in miRNAs and miRNA targets. *Nucleic Acids Res*. 2021;49(D1):D1276-1281.
183. Yue M, Zhou D, Zhi H, Wang P, Zhang Y, Gao Y, et al. MSDD: A manually curated database of experimentally supported associations among miRNAs, SNPs and human diseases. *Nucleic Acids Res*. 2018;46(D1):D181–5.

184. Bhattacharya A, Ziebarth JD, Cui Y. SomamiR: A database for somatic mutations impacting microRNA function in cancer. *Nucleic Acids Res.* 2013;41(Database issue):D977-982.
185. Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Ridwan Amode M, et al. Ensembl 2021. *Nucleic Acids Res.* 2021;49(D1):D884-891.
186. Consortium Gte. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science.* 2015;348(6235):648–60.
187. Ward LD, Kellis M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 2012;40(D1):D930-934.
188. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Res.* 2013;41(D1):D991-995.
189. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 2014;42(Database issue):D92-97.
190. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68–74.
191. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44–57.
192. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol.* 2020;38(6):675–8.
193. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 2022;48(D1):D20–6.
194. Kheradpour P, Kellis M. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Res.* 2014;42(5):2976–87.

195. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003;31(13):3406–15.
196. Wei CY, Tan QX, Zhu X, Qin QH, Zhu FB, Mo QG, et al. Expression of CDKN1A/p21 and TGFBR2 in breast cancer and their prognostic significance. *Int J Clin Exp Pathol.* 2015;8(11):14619–29.
197. Wang F, Ma YL, Zhang P, Yang JJ, Chen HQ, Liu ZH, et al. A genetic variant in microRNA-196a2 is associated with increased cancer risk: A meta-analysis. *Mol Biol Rep.* 2012;39(1):269–75.
198. Wang PY, Gao ZH, Jiang ZH, Li XX, Jiang BF, Xie SY. The associations of single nucleotide polymorphisms in miR-146a, miR-196a and miR-499 with breast cancer susceptibility. *PLoS One.* 2013;8(9):e70656.
199. Zhang H, Su YL, Yu H, Qian BY. Meta-analysis of the association between miR-196a-2 polymorphism and cancer susceptibility. *Cancer Biol Med.* 2012;9(1):63–72.
200. Hu Z, Liang J, Wang Z, Tian T, Zhou X, Chen J, et al. Common genetic variants in pre-microRNAs were associated with increased risk of breast cancer in Chinese women. *Hum Mutat.* 2009;30(1):79–84.
201. Zhao H, Xu J, Zhao D, Geng M, Ge H, Fu L, et al. Somatic mutation of the SNP rs11614913 and its association with increased MIR 196A2 expression in breast cancer. *DNA Cell Biol.* 2016;35(2):81–7.
202. Chen XY, Zhang J, Zhu JS. The role of m6A RNA methylation in human cancer. *Mol Cancer.* 2019;18(1):103.
203. Sun T, Wu R, Ming L. The role of m6A RNA methylation in cancer. *Biomed Pharmacother.* 2019;112:108613.
204. Lan Q, Liu PY, Haase J, Bell JL, Hüttelmaier S, Liu T. The critical role of RNA m6A methylation in cancer. *Cancer Res.* 2019;79(7):1285–92.
205. Vu LP, Cheng Y, Kharas MG. The biology of m6A RNA methylation in normal and malignant hematopoiesis. *Cancer Discov.* 2019;9(1):25–33.
206. Zhao W, Qi X, Liu L, Ma S, Liu J, Wu J. Epigenetic regulation of m6A modifications in human cancer. *Mol Ther Nucleic Acids.* 2020;19:405–12.

207. Li J, Chen Z, Su K, Zeng J. Clinicopathological classification and traditional prognostic indicators of breast cancer. *Int J Clin Exp Pathol.* 2015;8(7):8500–5.
208. Meyer KD, Jaffrey SR. Rethinking m6A readers, writers, and erasers. *Annu Rev Cell Dev Biol.* 2017;33:319–42.
209. Deng YN, Xia Z, Zhang P, Ejaz S, Liang S. Transcription factor RREB1: From target genes towards biological functions. *Int J Biol Sci.* 2020;16(8):1463–73.
210. Choupani J, Nariman-Saleh-Fam Z, Saadatian Z, Ouladsahebmadarek E, Masotti A, Bastami M. Association of mir-196a-2 rs11614913 and mir-149 rs2292832 polymorphisms with risk of cancer: An updated meta-analysis. *Front Genet.* 2019;10:186.
211. Catucci I, Yang R, Verderio P, Pizzamiglio S, Heesen L, Hemminki K, et al. Evaluation of SNPs in miR-146a, miR196a2 and miR-499 as low-penetrance alleles in German and Italian familial breast cancer cases. *Hum Mutat.* 2010;31(1):E1052-1057.
212. Jedlinski DJ, Gabrovska PN, Weinstein SR, Smith RA, Griffiths LR. Single nucleotide polymorphism in hsa-mir-196a-2 and breast cancer risk: A case control study. *Twin Res Hum Genet.* 2011;14(5):417–21.
213. Mousses S, özçelik H, D.lee P, Malkin D, B.bull S, L.andrulis I. Two variants of the CIP1/WAF1 gene occur together and are associated with human cancer. *Hum Mol Genet.* 1995;4(6):1089–92.
214. Bae I, Fan S, Bhatia K, Kohn KW, Fornace AJ, O'Connor PM. Relationships between G1 arrest and atability of the p53 and p21Cip1/Waf1 proteins following gamma-irradiation of human lymphoma cells. *Cancer Res.* 1995;55(11):2387–93.
215. Zhang Y, Fujita N, Tsuruo T. Caspase-mediated cleavage of p21Waf1/Cip1 converts cancer cells from growth arrest to undergoing apoptosis. *Oncogene.* 1999;18(5):1131–8.
216. Abbas T, Dutta A. P21 in cancer: Intricate networks and multiple activities. *Nat Rev Cancer.* 2010;9(6):400–14.
217. Kreis N-N, Louwen F, Yuan J. The multifaceted p21 (Cip1/Waf1/CDKN1A) in cell differentiation, migration and cancer therapy. *Cancers (Basel).* 2019;11(9):1220.

218. Carvalho INSR, de Oliveira Reis AH, Cabello PH, Vargas FR. Polymorphisms of CDKN1A gene and risk of retinoblastoma. *Carcinogenesis*. 2013;34(12):2774–7.
219. Lotfi Garavand A, Mohammadi M, Mohammadzadeh S. Evaluation of TP53 codon 72, P21 codon 31, and MDM2 SNP309 polymorphisms in Iranian patients with acute lymphocytic leukemia. *Reports Biochem Mol Biol*. 2020;9(1):26–32.
220. Lin YC, Hour TC, Tsai YC, Huang SP, Wu WJ, Chen CH, et al. Preliminary evidence of polymorphisms of cell cycle regulatory genes and their roles in urinary tract urothelial cancer susceptibility and prognosis in a Taiwan population. *Urol Oncol Semin Orig Investig*. 2017;35(9):543.e7-16.
221. Sjölander A, Birgander R, Rannug A, Alexandrie AK, Tornling G, Beckman G. Association between the p21 codon 31 A1 (arg) allele and lung cancer. *Hum Hered*. 1996;46(4):221–5.
222. Li G, Liu Z, Sturgis EM, Shi Q, Chamberlain RM, Spitz MR, et al. Genetic polymorphisms of p21 are associated with risk of squamous cell carcinoma of the head and neck. *Carcinogenesis*. 2005;26(9):1596–602.
223. Barbieri RB, Bufalo NE, Secolin R, Assumpção LVM, Maciel RMB, Cerutti JM, et al. Polymorphisms of cell cycle control genes influence the development of sporadic medullary thyroid carcinoma. *Eur J Endocrinol*. 2014;171(6):761–7.
224. Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MWR, Pooley KA, et al. A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet*. 2007;39(3):352–8.
225. Driver KE, Song H, Lesueur F, Ahmed S, Barbosa-Morais NL, Tyrer JP, et al. Association of single-nucleotide polymorphisms in the cell cycle genes with breast cancer in the British population. *Carcinogenesis*. 2008;29(2):333–41.
226. Machnik M, Cylwa R, Kiełczewski K, Biecek P, Liloglou T, Mackiewicz A, et al. The expression signature of cancer-associated KRAB-ZNF factors identified in TCGA pan-cancer transcriptomic data. *Mol Oncol*. 2019;13(4):701–24.

227. Liu L, Lei X, Meng J, Wei Z. WITMSG: Large-scale prediction of human intronic m6A RNA methylation sites from sequence and genomic features. *Curr Genomics*. 2020;21(1):67–76.
228. Yee LD, Mortimer JE, Natarajan R, Dietze EC, Seewaldt VL. Metabolic health, insulin, and breast cancer: Why oncologists should care about insulin. *Front Endocrinol (Lausanne)*. 2020;11:58.
229. Goodwin JP. Insulin resistance in breast cancer: Relevance and clinical implications. *Breast Cancer Res*. 2011;13(Suppl 2):O7.
230. Shen T, Lu Y, Zhang Q. High squalene epoxidase in tumors predicts worse survival in patients with hepatocellular carcinoma: Integrated bioinformatic analysis on NAFLD and HCC. *Cancer Control*. 2020;27:1–9.
231. Lee H, Kim K, Woo J, Park J, Kim H, Lee KE, et al. Quantitative proteomic analysis identifies AHNAK (Neuroblast Differentiation-associated Protein AHNAK) as a novel candidate biomarker for bladder urothelial carcinoma diagnosis by liquid-based cytology. *Mol Cell Proteomics*. 2018;17(9):1788–802.
232. Ma D, Pan Z, Chang Q, Zhang J jin, Liu X, Hua N, et al. KLF5-mediated Eppk1 expression promotes cell proliferation in cervical cancer via the p38 signaling pathway. *BMC Cancer*. 2021;21:377.
233. Guo X, Hao Y, Kamilijiang M, Hasimu A, Yuan J, Wu G, et al. Potential predictive plasma biomarkers for cervical cancer by 2D-DIGE proteomics and ingenuity pathway analysis. *Tumor Biol*. 2015;36(3):1711–20.
234. Smith A, Poole L, Dhanwada K, Goonesekere N. Identification of candidate biomarkers and cancer genes AHNAK2 and EPPK1 in pancreatic cancer. *Br J Med Med Res*. 2016;18(8):1–8.
235. Jing-Sheng Z, Zhen-Don Gz, Li P, Zhi-Zhu B, Yong Y, Hong Y, et al. CBX4 exhibits oncogenic activities in breast cancer via Notch1 signaling. *Int J Biochem Cell Biol*. 2018;95:1–8.
236. Akkiprik M, Peker İ, Özmen T, Amuran GG, Güllüoğlu BM, Kaya H, et al. Identification of differentially expressed IGFBP5-related genes in breast cancer tumor tissues using cDNA microarray experiments. *Genes (Basel)*. 2015;6(4):1201–14.

237. Karabulut S, Kaya Z, Amuran GG, Peker I, Özmen T, Gulluolu BM, et al. Correlation between the DNA methylation and gene expression of IGFBP5 in breast cancer. *Breast Dis.* 2016;36(4):123–31.
238. Sureshbabu A, Okajima H, Yamanaka D, Tonner E, Shastri S, Maycock J, et al. IGFBP5 induces cell adhesion, increases cell survival and inhibits cell migration in MCF-7 human breast cancer cells. *J Cell Sci.* 2012;125(7):1693–705.
239. Akkiprik M, Hu L, Sahin A, Hao X, Zhang W. The subcellular localization of IGFBP5 affects its cell growth and migration functions in breast cancer. *BMC Cancer.* 2009;9:103.
240. Mangone FR, Valoyes MAV, Do Nascimento RG, Conceição MPF, Bastos DR, Pavanelli AC, et al. Prognostic and predictive value of Pleckstrin homology-like domain, family a family members in breast cancer. *Biomark Med.* 2020;14(16):1537–52.
241. Johnson EO, Chang KH, de Pablo Y, Ghosh S, Mehta R, Badve S, et al. PHLDA1 is a crucial negative regulator and effector of Aurora A kinase in breast cancer. *J Cell Sci.* 2011;124(16):2711–22.
242. Nagai MA, Fregnani JHTG, Netto MM, Brentani MM, Soares FA. Down-regulation of PHLDA1 gene expression is associated with breast cancer progression. *Breast Cancer Res Treat.* 2007;106:49–56.
243. Taylor S, Lam M, Pararasa C, Brown JEP, Carmichael AR, Griffiths HR. Evaluating the evidence for targeting FOXO3a in breast cancer: A systematic review. *Cancer Cell Int.* 2015;15:1.
244. Alday-Parejo B, Richard F, Wörthmüller J, Rau T, Galván JA, Desmedt C, et al. MAGI1, a new potential tumor suppressor gene in estrogen receptor positive breast cancer. *Cancers (Basel).* 2020;12:223.
245. Yoe J, Kim D, Kim S, Lee Y. Capicua restricts cancer stem cell-like properties in breast cancer cells. *Oncogene.* 2020;39(17):3489–506.
246. Schulte I, Batty EM, Pole JCM, Blood KA, Mo S, Cooke SL, et al. Structural analysis of the genome of breast cancer cell line ZR-75-30 identifies twelve expressed fusion genes. *BMC Genomics.* 2012;13:719.

La N⁶-metiladenosina (m⁶A) es una de las modificaciones más abundantes del ARNm presente en los organismos eucariotas. Distintos estudios han revelado que la modificación m⁶A está involucrada en la regulación de la mayoría de los procesos del metabolismo del ARNm. Sin embargo, aún no existe mucho conocimiento acerca del impacto que esta modificación tiene sobre la regulación de la expresión génica, su función en los procesos metabólicos y el efecto que puede ejercer en el desarrollo y progresión de ciertas enfermedades, como, por ejemplo, el cáncer de mama. Por tal motivo, en este estudio, exploramos la presencia de la modificación m⁶A en el cáncer de mama, específicamente en los subtipos moleculares luminal A, luminal B HER2-, luminal B HER2+, HER2+ y triple negativo, con el fin de determinar el estado de la metilación del ARN y de su posible relación con futuras dianas terapéuticas. Para ello, se ha evaluado el estado de la metilación m⁶A y de la expresión de las principales metiltransferasas y desmetilasas en muestras provenientes de los mencionados subtipos moleculares de cáncer de mama. También, se han identificado m⁶A-SNPs relacionados con esta enfermedad y se ha determinado el perfil de metilación m⁶A presente en todo el transcriptoma de los distintos subtipos moleculares mediante el uso de la técnica MeRIP-seq.