



ESCUELA DE DOCTORADO
INTERNACIONAL DE LA USC

Vera
García Outeiral

Tesis doctoral

DNMT2 y plasticidad celular:
análisis de sus funciones en
pluripotencia, reprogramación
celular y cáncer

Santiago de Compostela, 2023

Programa de doctorado en Medicina Molecular

TESIS DOCTORAL

DNMT2 Y PLASTICIDAD CELULAR: ANÁLISIS DE SUS FUNCIONES EN PLURIPOTENCIA, REPROGRAMACIÓN CELULAR Y CÁNCER

Autor

Vera García Outeiral

Directores: Juan Bautista Zalvide Torrente y Miguel Ángel Fidalgo Pérez

Tutor: Miguel Ángel Fidalgo Pérez

PROGRAMA DE DOCTORADO EN MEDICINA MOLECULAR

SANTIAGO DE COMPOSTELA

DECLARACIÓN DE CONFLICTO DE INTERESES Y AUTORÍA DE LOS DATOS PRESENTADOS

Vera García Outeiral declara no tener ningún conflicto de interés en relación con la presente tesis doctoral.

Por la presente, la doctoranda declara la completa autoría de todas las imágenes y figuras mostradas en la tesis.

En Santiago de Compostela, 14 de diciembre de 2023

Durante el desarrollo de esta tesis doctoral, la investigación en el grupo del Dr. Miguel Fidalgo ha sido respaldada financieramente por la Agencia Estatal de Investigación (FEDER/Ministerio de Ciencia, Innovación y Universidades–AEI/BFU2016-80899-P y PID2019-105739GB-I00/AEI/10.13039/501100011033) y una ayuda del programa Ramón y Cajal (RYC-2014-16779) del Ministerio de Economía y Competitividad de España. A lo largo de este período, Vera García Outeiral fue beneficiaria de una ayuda para la Formación del Profesorado Universitario (FPU17/01131), otorgada por el Ministerio de Educación, Cultura y Deporte. Además, recibió apoyo a través de una ayuda a la movilidad para estancias breves y traslados temporales de beneficiarios FPU (EST22/00548) del Ministerio de Universidades.

AGRADECIMIENTOS

Considero que la felicidad que no se comparte tiene un valor limitado. Por esta razón, quiero referirme aquí a aquellos cuyo apoyo ha sido fundamental para mí durante esta etapa.

En primer lugar, quiero agradecer a mis directores de tesis, el Dr. Juan Zalvide y el Dr. Miguel Fidalgo, por haberme dado la oportunidad de realizar esta tesis doctoral. En especial a mi tutor, Miguel, por haberme acogido en el laboratorio *Stem Cells & Human Diseases* y por la excepcional formación que he recibido todos estos años. Gracias por motivarme a dar lo mejor de mí y por promover mi pensamiento crítico, mi ambición y mi desarrollo como investigadora. También quiero agradecer a la Dra. Diana Guallar, por valorar mis habilidades científicas y por sus valiosos consejos para afrontar el ámbito académico.

Como no, debo mencionar a la primera generación de doctores del laboratorio *Stem Cells & Human Diseases*. A Parda, por hacerme sentir siempre como en casa; a Álex, por darme, literalmente, el empujón que necesitaba para unirme al laboratorio; y sobre todo, a Yara, quien además de una compañera de trabajo impecable, fue y sigue siendo una gran amiga fuera de aquí. Aunque nos separan miles de kilómetros de distancia, sigues estando presente. Además, mi agradecimiento a Cris A, por su constante apoyo en este camino y por brindarme siempre su confianza. También a los miembros, actuales y pasados, de la unidad *Stem Cells & Epitranscriptomics*, dirigida por el Dr. Miguel Fidalgo y la Dra. Diana Guallar. Especialmente a Tiago, Cris DP, Adrián, Helena, Alba y David. Gracias por los buenos momentos dentro y fuera del laboratorio. Está claro que han hecho el recorrido mucho más fácil.

I also want to thank Dr. Carlos-Filipe Pereira for the exceptional opportunity provided to me to undertake a predoctoral stay at the *Cell Reprogramming and Immunity Lab* at Lund University. The experience has been profoundly enlightening, offering me invaluable insights and fostering personal and academic growth. I extend my deepest gratitude to the entire team, with special mention to Rita and Evelyn, whose unwavering support and friendship became indispensable pillars during my time there.

Si tuviera que citar a uno de los apoyos más fundamentales de este viaje sería a Adriana. Gracias por demostrarme la autenticidad de actuar con el corazón, por apoyarme inquebrantablemente todos estos años y por ser la mejor compañera y amiga que una podría querer a su lado. Aquí también quiero mencionar a Lucía, otra parte imprescindible de la ecuación. Gracias por confiar firmemente en mí y en mis capacidades. Tu comprensión incondicional significa muchísimo para mí y me alegro profundamente de habernos cruzado en este camino. Por supuesto, sé que voy a llevaros a ambas siempre conmigo.

Evidentemente, entre las personas especiales quiero destacar a Tomás. Gracias por conseguir que mantenga mis pies en la tierra y mi cabeza en las nubes al mismo tiempo. Todo tú me hace explotar de orgullo y admiración. Estoy deseando ver qué nos depara esta nueva aventura.

No puedo dejar de hacer referencia a mi grupo de biología. En especial a Toño, Alberto, Pablito, Iván, Emilio, Iria, Zoe, Andrea, Yi y Mery, por hacerme sentir la estrella que más brilla de nuestro microuniverso. Tampoco a mi grupo del CiMUS, por su poder revitalizador y especialmente a Sarru, Joan, Dami y Leo, por acogerme y hacerme sentir siempre como una más del núcleo. No olvido a los miembros de “Lund reducido”, particularmente a Ángeles y Nacho. Os doy las gracias por las comidas en Velvet, los viajes a Copenhague y Gotemburgo y

por todos los momentos que hemos compartido juntos. Habéis hecho de Lund un segundo hogar para mí.

Quiero mencionar también a mis amigas de toda la vida: Lorena, Laura, Carme, Alicia y Xeila. Somos muy distintas, pero nos une una química especial. Pienso en nosotras como ejemplos de perseverancia, ambición e inquietud y no puedo estar más orgullosa de cada una de vosotras.

Finalmente, tengo que agradecer todo lo logrado a mi familia. A mis tíos, Carmela, Barral y Carlos. Ojalá algún día verme con vuestros ojos. A mis abuelos, que aunque ya no están, siguen haciéndome sentir infinitamente especial. A mi hermana, Sofía, por su apoyo incondicional, y sin lugar a duda, a mis padres. Gracias por darlo todo por mí. Por alentar mi curiosidad desde la cuna y por enseñarme que el esfuerzo finalmente tiene su recompensa. Vuestra paciencia infinita y creencia en mí han sido el mayor motor a lo largo de este viaje. Sois mi gran ejemplo de determinación y habéis sido el pilar más estable en cada paso que he dado. Os quiero.

A mis padres,

CONTENIDO

1. ABSTRACT.....	3
2. RESUMEN.....	7
3. INTRODUCCIÓN.....	11
3.1 El estado celular pluripotente.....	13
3.1.1 Las células troncales embrionarias	13
3.1.2 La regulación del estado celular pluripotente	14
3.2 La pluripotencia inducida	15
3.2.1 Inducción de pluripotencia mediada por OSKM	15
3.2.2 Etapas del proceso de reprogramación celular	17
3.2.2.1. Fase de iniciación.....	18
3.2.2.2. Fases de maduración y estabilización	19
3.3 Modificaciones químicas en ARN	20
3.3.1 Visión general	20
3.3.2 Distribución, funciones y regulación de la marca m ⁵ C.....	21
3.4 DNMT2.....	26
3.4.1 Filogenia, estructura y actividad catalítica	26
3.4.2 Funciones biológicas de DNMT2	29
3.4.3 Fenotipo y procesos patológicos asociados	33
4. OBJETIVOS.....	37
5. MATERIALES Y MÉTODOS.....	41
5.1 Reactivos y kits	43
5.2 Modelos experimentales.....	43
5.3 Cultivo celular	44
5.4 Generación de vectores plasmídicos	45
5.5 Transfección de células de mamífero	50
5.6 Generación de ESCs KO y KI mediante CRISPR-Cas9	51
5.7 Producción de virus y transducción celular	52
5.8 Extracción y purificación de ADN genómico.....	53
5.9 Reacción en cadena de la polimerasa.....	53
5.10 Extracción y purificación de ARN.....	54
5.11 Obtención de ADN complementario.....	54
5.12 Reacción en cadena de la polimerasa cuantitativa con transcripción inversa	55
5.13 Extracción y purificación de proteína total.....	55
5.14 Western blot.....	55
5.15 Ensayos de reprogramación celular	56
5.16 Ensayos de diferenciación	58
5.17 Ensayos de formación de colonias.....	60
5.18 Tinciones	61
5.19 Inmunofluorescencia	62

5.20	Análisis de ciclo celular	62
5.21	Citometría de flujo.....	63
5.22	Ensayos de proliferación	63
5.23	Ensayo de cierre de heridas.....	64
5.24	Determinación de m ⁵ C en ARN	64
5.25	Secuenciación masiva de ARN y análisis	65
5.26	Análisis <i>in silico</i> de los datos transcriptómicos	66
5.27	Identificación del interactoma de DNMT2 en ESCs	68
5.28	Obtención y análisis de datos de pacientes con cáncer	70
5.29	Análisis filogenéticos.....	70
5.30	Análisis estadísticos	70
5.31	Recursos bibliográficos	71
5.32	Programas y algoritmos.....	71
6.	RESULTADOS	73
6.1	Análisis del papel de DNMT2 en la reprogramación celular somática	75
6.1.1	La presencia de DNMT2 es necesaria para la generación de iPSCs	75
6.1.2	La ausencia de DNMT2 impide la MET durante reprogramación	77
6.1.3	La función catalítica de DNMT2 es necesaria durante la MET	82
6.1.4	La función de DNMT2 es prescindible en etapas avanzadas de la reprogramación celular somática	85
6.1.5	DNMT2 interviene en la reprogramación celular directa a precursores hematopoyéticos.....	86
6.2	Determinación del papel de DNMT2 en la regulación del estado celular pluripotente	91
6.2.1	La ausencia de DNMT2 es compatible con el mantenimiento de pluripotencia	91
6.2.2	La ausencia de DNMT2 altera el transcriptoma de las ESCs	98
6.2.3	La función de DNMT2 es necesaria para la salida de pluripotencia	104
6.2.4	La función m ⁵ C-metiltransferasa de DNMT2 en el ARN es necesaria en pluripotencia	109
6.2.5	Identificación de los interactores de DNMT2 en ESCs	118
6.3	Exploración del papel de DNMT2 en cáncer	121
6.3.1	Las mutaciones en el gen <i>DNMT2</i> o la reducción de sus niveles de expresión se asocian con un peor pronóstico en varios tipos de cáncer	121
6.3.2	El silenciamiento de <i>DNMT2</i> aumenta la agresividad de la línea de cáncer de mama MCF-7	123
6.3.3	DNMT2 participa en la regulación de P53 en la línea tumoral MCF-7	128
7.	Discusión	131
7.1	La contribución de DNMT2 a la reprogramación celular	133
7.2	La relevancia de DNMT2 en pluripotencia	137
7.3	El papel de DNMT2 en cáncer de mama	142
8.	CONCLUSIONS / CONCLUSIONES	145
9.	APÉNDICE.....	149
10.	LISTA DE ABREVIATURAS	159
11.	LISTA DE FIGURAS	165
12.	BIBLIOGRAFÍA	171

1. ABSTRACT

Pluripotent cells, including ESCs and iPSCs, harbor immense therapeutic potential due to their capacity for indefinitely self-renewal in culture and their ability to differentiate into any cell type within the adult organism. Additionally, they serve as an effective platform to investigate the molecular foundations controlling cellular plasticity. Understanding the intricate regulatory networks governing pluripotency, along with the similarities between pluripotency and cancer, has propelled research focused on analyzing regulatory factors, such as DNA methyltransferases (DNMTs), in the inherent plasticity of pluripotent and cancer cells. However, the role of DNMT2, recognized as the most evolutionarily conserved and enigmatic member of this family of proteins, due to its exclusive catalytic activity in RNA rather than DNA in mammalian organisms, remains unclear in the context of pluripotency and cancer. Motivated by this, in this doctoral thesis, we aimed to investigate the dependent and independent functions of DNMT2's catalytic activity in cellular reprogramming processes, including the acquisition and exit from the pluripotent state, as well as in the biology of pluripotent and cancer cells. To this end, we employed loss and gain-of-function models of DNMT2, dependent and independent of its 5-methylcytosine (m^5C)-methyltransferase activity in RNA, using a wide array of cellular and biochemical approaches, omics techniques, and bioinformatic analyses. The results underscore the essential role of DNMT2's catalytic activity in RNA during cellular reprogramming for the acquisition and exit from the pluripotent state. Although DNMT2 was found to be dispensable for maintaining pluripotency, the presence of mutations or decreased expression was associated with an unfavorable prognosis in various types of cancer, including breast cancer. Consistently, silencing DNMT2 intensified the aggressiveness of a luminal A subtype associated breast cancer cell line. These findings unveil a new regulatory layer, mediated by the chemical modification of RNA, m^5C , through DNMT2, dedicated to reprogramming and pluripotency. Furthermore, they open new avenues to explore the dependent and independent functions of DNMT2's catalytic activity in cellular processes related to cancer.

2. RESUMEN

Las células pluripotentes, que incluyen a las ESCs y a las iPSCs, presentan un gran potencial terapéutico debido a su capacidad de autorrenovarse indefinidamente en cultivo y de diferenciarse a cualquier tipo celular del organismo adulto. Además, constituyen una eficaz plataforma para investigar los fundamentos moleculares que controlan la plasticidad celular. La comprensión de las intrincadas redes reguladoras que gobiernan la pluripotencia, junto con las similitudes entre esta y el cáncer, ha impulsado investigaciones centradas en analizar factores reguladores, como las enzimas metiltransferasas de ADN (DNMTs), en la plasticidad inherente de las células pluripotentes y las células cancerosas. No obstante, la función de DNMT2, el miembro más conservado y enigmático de esta familia de proteínas, debido a su actividad catalítica exclusiva en ARN en lugar de en ADN en mamíferos, permanece sin dilucidar en los contextos de pluripotencia y cáncer. Motivados por esto, en esta tesis doctoral nos propusimos investigar las funciones dependientes e independientes de la actividad catalítica de DNMT2 en los procesos de reprogramación celular, que incluyen la adquisición y salida del estado pluripotente, así como en la biología de células pluripotentes y cancerosas. Para este fin, se emplearon modelos de pérdida y ganancia de función de DNMT2, dependientes e independientes de su actividad 5-metilcitosina (m^5C)-metiltransferasa en ARN, con una amplia variedad de enfoques celulares y bioquímicos, técnicas ómicas y análisis bioinformáticos. Los resultados obtenidos resaltan la relevancia de la actividad catalítica de DNMT2 en ARN durante la reprogramación celular para la adquisición y para la salida del estado pluripotente. Aunque DNMT2 resultó ser prescindible para el mantenimiento de la pluripotencia, la presencia de mutaciones o la disminución de su expresión se asociaron con un pronóstico desfavorable en diversos tipos de cáncer, incluyendo el de mama. En consonancia, su silenciamiento intensificó la agresividad de una línea celular tumoral asociada al subtipo luminal A de cáncer de mama. Estos hallazgos revelan la existencia de un nuevo estrato regulador, mediado por la modificación m^5C del ARN dependiente de DNMT2, dedicado a la reprogramación celular y pluripotencia. Además, generan nuevas perspectivas para explorar las funciones dependientes e independientes de la actividad catalítica de DNMT2 en procesos celulares relacionados con el cáncer.

3. INTRODUCCIÓN

3.1 EL ESTADO CELULAR PLURIPOTENTE

3.1.1 Las células troncales embrionarias

La pluripotencia en mamíferos es una característica única y transitoria que presentan algunas células durante las etapas iniciales del desarrollo embrionario (Hassani et al., 2019; Kinoshita y Smith, 2018). Esto abarca desde las células que componen el epiblasto preimplantacional hasta aquellas del epiblasto postimplantacional en el útero materno. Implica la capacidad de generar los distintos tipos celulares de las tres capas germinales (ectodermo, mesodermo y endodermo), así como las células germinales primordiales, dando origen, en última instancia, a cada una de las células especializadas que conforman un organismo adulto. Dada su naturaleza temporal, este potencial de diferenciación se restringe gradualmente a medida que progresa el desarrollo (Hassani et al., 2019; Kinoshita y Smith, 2018).

A través de la derivación de algunas células de la masa celular interna (ICM, por sus siglas en inglés) del blastocisto preimplantacional, la pluripotencia puede ser capturada en cultivo, dando origen a las células troncales embrionarias (ESCs, por sus siglas en inglés) (Evans y Kaufman, 1981; Hassani et al., 2019; Martin, 1981; Thomson et al., 1998) (**Figura I1**). En contraste con sus homólogas *in vivo*, estas células tienen la capacidad única de autorrenovarse de manera indefinida bajo condiciones específicas de cultivo, al mismo tiempo que mantienen la habilidad de diferenciarse en cada uno de los tipos celulares altamente especializados que constituyen un individuo adulto (Hassani et al., 2019; Kinoshita y Smith, 2018). La derivación de ESCs ha posibilitado profundizar en la investigación de los mecanismos moleculares que regulan el estado celular pluripotente, así como los procesos de diferenciación celular durante la salida de este, los cuales están sometidos a una estricta regulación transcripcional, epigenética, metabólica y de diversas vías de señalización celular, entre otros (De Los Angeles et al., 2015; Kinoshita y Smith, 2018; Li y Izpisua Belmonte, 2018). Además, estas células representan una fuente prácticamente inagotable de recursos celulares, lo que fomenta su aplicación en la modelización de enfermedades y en el descubrimiento y evaluación de fármacos, presentando así un gran potencial en medicina regenerativa (Pendse et al., 2022; Yamanaka, 2020).

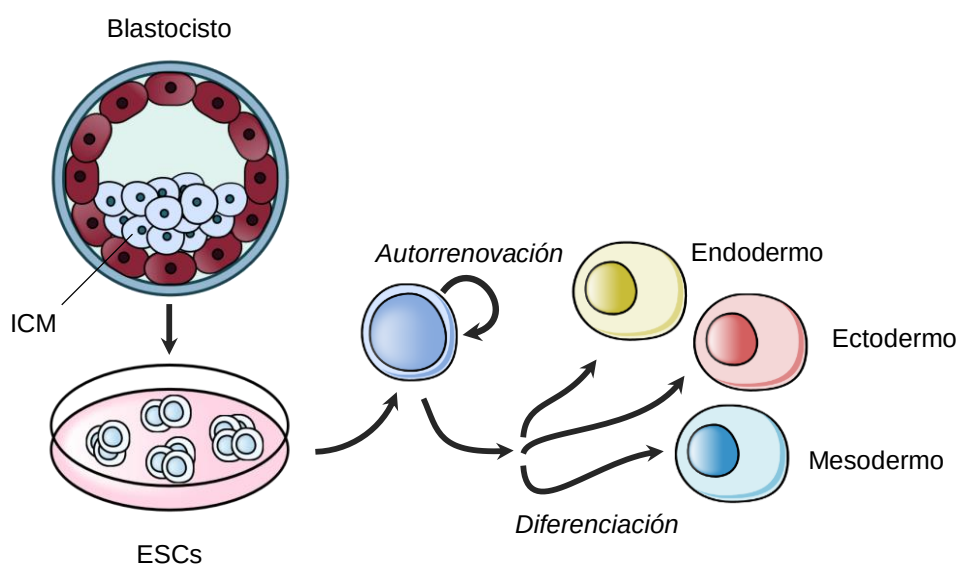


Figura I1. Características principales de las ESCs. Esta representación simplificada ilustra cómo las células troncales embrionarias (ESCs), derivadas de la masa celular interna (ICM) del blastocisto murino, tienen la

capacidad de autorrenovarse de manera ilimitada en cultivo. Además, muestran su habilidad para diferenciarse en células precursoras de cada una de las tres capas germinales: ectodermo, endodermo y mesodermo. Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

3.1.2 La regulación del estado celular pluripotente

Múltiples evidencias *in vivo* e *in vitro* han confirmado que el estado celular pluripotente se rige por una red reguladora de genes de pluripotencia (PGRN, por sus siglas en inglés), interconectada mediante un circuito central conformado por los factores de transcripción OCT4, SOX2 y NANOG (Kim et al., 2008; Li y Belmonte, 2017; Li y Izpisua Belmonte, 2018). La PGRN se caracteriza por su bivalencia, ya que, a diferencia de lo que ocurre en otros estados celulares, debe mantener la capacidad de propagación indefinida del estado pluripotente mediante la supresión activa de la especificación del linaje celular, a la vez que un potencial de diferenciación amplio e inesgado (Kim et al., 2008; Li y Belmonte, 2017; Li y Izpisua Belmonte, 2018).

Los factores de transcripción OCT4, SOX2 y NANOG desempeñan un papel fundamental tanto en el mantenimiento y salida de la pluripotencia (Masui et al., 2007; Mitsui et al., 2003; Nichols et al., 1998), como en su adquisición (Takahashi et al., 2007; Takahashi y Yamanaka, 2006; Yu et al., 2007). De hecho, la expresión forzada de combinaciones específicas de estos y otros factores de transcripción es suficiente para eludir los mecanismos moleculares que salvaguardan el destino celular somático (Brumbaugh et al., 2019). Esto resulta en el restablecimiento del estado celular pluripotente, dando origen a las células pluripotentes inducidas o iPSCs (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007), como se detallará en la siguiente sección. OCT4, SOX2 y NANOG actúan de manera cooperativa, a menudo sobre las mismas dianas genómicas, formando circuitos reguladores de retroalimentación que permiten el mantenimiento de su propia expresión, así como la de otros genes que codifican factores esenciales en la regulación del estado celular pluripotente (Boyer et al., 2005; Chen et al., 2008; Chew et al., 2005; Loh et al., 2006; Masui et al., 2007; Rodda et al., 2005).

OCT4 se encuentra principalmente en células pluripotentes, siendo esencial tanto para el mantenimiento de la pluripotencia *in vivo* como en cultivo celular (Nichols et al., 1998). Asimismo, la expresión de SOX2 es necesaria durante el desarrollo embrionario para la formación del epiblasto pluripotente contenido en la ICM del blastocisto preimplantacional (Avilion et al., 2003; Thomson et al., 2011). Además, su activación representa un paso limitante durante la adquisición de pluripotencia mediante reprogramación celular somática (Buganim et al., 2012). Teniendo esto en cuenta, la regulación precisa de ambos factores es crucial para un correcto mantenimiento de la PGRN y por ende, del estado celular pluripotente. Es más, la pérdida de OCT4 o SOX2 promueve la diferenciación hacia el linaje extraembrionario trofectodermo (Fong et al., 2008; Masui et al., 2007; Niwa et al., 2000), mientras que la ganancia de función de OCT4 o SOX2 conduce a la diferenciación hacia mesendodermo o ectodermo neural, respectivamente (Thomson et al., 2011). Por otro lado, NANOG también es un factor central de pluripotencia. Esto se debe a su importante papel durante el desarrollo embrionario temprano para la formación del epiblasto pluripotente presente en células de la ICM (Chambers et al., 2003; Mitsui et al., 2003). Aunque su ausencia es compatible con el mantenimiento de las ESCs bajo condiciones específicas de cultivo celular, la expresión constitutiva de NANOG facilita la autorrenovación de las ESCs en cultivo de manera independiente de la presencia del factor inhibidor de la leucemia (LIF, por sus siglas en inglés),

una citocina clave durante este proceso (Chambers et al., 2003; Chambers et al., 2007; Mitsui et al., 2003).

Conjuntamente, los factores centrales de pluripotencia proporcionan una plataforma para el reclutamiento de factores adicionales en la cromatina que integran señales reguladoras procedentes de distintas rutas, así como otros circuitos transcripcionales (Kim et al., 2008; Li y Belmonte, 2017; Li y Izpisua Belmonte, 2018). Además del control transcripcional, la PGRN también se regula por otros mecanismos que incluyen: el procesamiento postranscripcional del ARN, la modificación de proteínas, la regulación metabólica y las modificaciones epigenéticas y epitranscriptómicas (**Figura I2**) (Che et al., 2022; Kim et al., 2008; Li y Belmonte, 2017; Li y Izpisua Belmonte, 2018; Liu et al., 2020c).

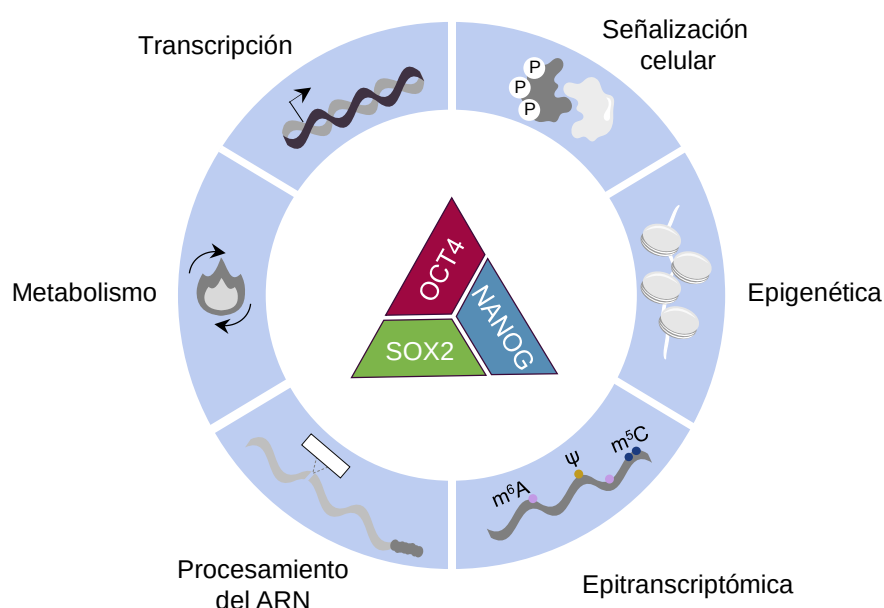


Figura I2. Mecanismos reguladores del estado celular pluripotente. Representación esquemática de los principales mecanismos que participan en la regulación de la pluripotencia, en combinación con el circuito central de la PGRN conformado por las proteínas OCT4, SOX2 y NANOG. Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

3.2 LA PLURIPOTENCIA INDUCIDA

3.2.1 Inducción de pluripotencia mediada por OSKM

La reprogramación celular somática (SCR, por sus siglas en inglés), constituye una vía alternativa para la obtención de células pluripotentes a partir de células ya diferenciadas (Shi et al., 2017; Yamanaka y Blau, 2010). La generación de células pluripotentes puede llevarse a cabo mediante aproximaciones muy distintas, como la transferencia nuclear de células somáticas a un óvulo enucleado (SCNT, por sus siglas en inglés) (Byrne et al., 2007; Wakayama et al., 2001), la fusión con células pluripotentes (Tada et al., 2001; Ying et al., 2002) o la expresión forzada en células somáticas de factores de transcripción específicos (Takahashi et al., 2007; Takahashi y Yamanaka, 2006). Aunque los métodos de fusión celular y SCNT son más rápidos, su investigación es dificultada por el gran número de productos génicos no identificados que median el proceso de reprogramación (Buganim et al., 2013; Takahashi y Yamanaka, 2015). Este inconveniente se solventa en el caso de la reprogramación mediada por

factores de transcripción, ya que se conocen los factores que inician el proceso, lo que permite una mejor modulación y seguimiento del mismo (Buganim et al., 2013).

La reprogramación celular mediante combinación de factores definidos no solo ha sido empleada para la obtención de células pluripotentes, sino también para convertir células ya diferenciadas en otros tipos celulares, sin necesidad de inducir un estado pluripotente intermedio, proceso conocido como reprogramación celular directa o transdiferenciación (Wang et al., 2021).

En 2006, Kazutoshi Takahashi y Shinya Yamanaka demostraron por primera vez que la expresión ectópica de solo cuatro factores de transcripción (OCT4, SOX2, KLF4 y cMYC, conocidos como OSKM), permitía reprogramar células somáticas, específicamente fibroblastos de ratón, hacia células con características pluripotentes, denominadas células troncales pluripotentes inducidas (iPSCs) (Takahashi y Yamanaka, 2006) (**Figura I3**). Estas iPSCs, prácticamente idénticas a las ESCs, tanto a nivel transcripcional como funcional, exhiben la capacidad de renovarse indefinidamente *in vitro* y de diferenciarse en cada uno de los tipos celulares que contiene un individuo adulto (Hochedlinger y Jaenisch, 2015; Shi et al., 2017).

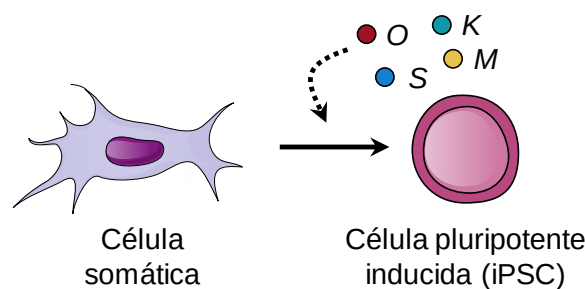


Figura I3. Reprogramación celular somática mediada por OSKM. La transducción de células somáticas con los factores de transcripción OCT4 (O), SOX2 (S), KLF4 (K) y cMYC (M), en conjunto referidos como OSKM, conduce a la generación de células pluripotentes inducidas (iPSCs). Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

Desde este descubrimiento histórico, que fue reconocido con el Premio Nobel de Medicina o Fisiología en 2012 al Dr. Yamanaka, ha habido un notable avance en las aproximaciones para obtener iPSCs. Estos progresos han abarcado diversas fuentes celulares somáticas de una amplia variedad de mamíferos, incluyendo al ser humano (Takahashi et al., 2007). Asimismo, se han desarrollado diversas estrategias para la generación de iPSCs, haciendo uso de variadas combinaciones de factores de transcripción, reguladores epigenéticos y pequeñas moléculas, entre otros (Hochedlinger y Jaenisch, 2015; Mao et al., 2022; Shi et al., 2017).

La relevancia de esta tecnología radica en que ofrece una alternativa a las ESCs, superando los problemas éticos asociados con la destrucción de embriones para su obtención (Fatima et al., 2023; Pendse et al., 2022). Además, la reprogramación celular ha abierto nuevas posibilidades en el modelado de enfermedades, así como en la investigación y descubrimiento de fármacos, detección de toxicidad y desarrollo de enfoques terapéuticos innovadores, tanto en el ámbito de la medicina general como en la personalizada (Fatima et al., 2023; Mao et al., 2022; Pendse et al., 2022; Rowe y Daley, 2019; Yamanaka, 2020).

A pesar de los avances logrados en la tecnología de SCR para la obtención de iPSCs, todavía existen desafíos asociados que requieren atención. Entre los problemas más frecuentes

se incluyen: la baja eficiencia del proceso, la inestabilidad genómica, la retención de la memoria epigenética, el riesgo de transformación oncogénica, el rechazo inmunológico, la heterogeneidad y la falta de madurez en las células después de la diferenciación (Fatima et al., 2023; Mao et al., 2022; Pendse et al., 2022; Yamanaka, 2020). Estos obstáculos deben ser abordados de manera integral para desbloquear el potencial completo de esta prometedora tecnología con fines terapéuticos.

Aunque han pasado casi dos décadas desde su descubrimiento inicial, aún desconocemos en detalle los mecanismos moleculares que orquestan este complejo proceso de SCR hacia células pluripotentes. Por lo tanto, la comprensión detallada de sus mecanismos reguladores es esencial para superar los desafíos actuales y permitir el desarrollo seguro y eficaz de las aplicaciones terapéuticas basadas en iPSCs.

Por otro lado, la conversión de células diferenciadas en iPSCs mediante la expresión de OSKM es ampliamente reconocida como una herramienta poderosa, similar a las ESCs, para comprender los mecanismos moleculares que controlan la plasticidad celular, que permite que las células asuman cambios en su destino o funciones, lo que continúa siendo uno de los principales desafíos de la biología moderna.

3.2.2 Etapas del proceso de reprogramación celular

La obtención de iPSCs es un proceso largo y complejo que tiene lugar mediante una secuencia ordenada de eventos, cada uno de los cuales actúa como un punto crítico limitante en el número de células que logran completar con éxito la etapa (Apostolou y Stadtfeld, 2018; Deng et al., 2021). Bajo condiciones estándar de cultivo, es decir, en presencia de suero y LIF, el proceso de SCR suele ser lento e ineficiente. Se observan trayectorias celulares distintas, que incluyen células "permisivas" (intermediarios de inicio, maduración y estabilización), que experimentan una reprogramación exitosa, y células "refractarias" (células con identidad estromal o similares a trofoblasto, endodermo extraembrionario, neuronas), que no logran formar iPSCs (Deng et al., 2021).

A continuación, con el propósito de facilitar la comprensión de los resultados obtenidos en esta tesis doctoral en el contexto de la reprogramación mediante los factores OSKM para generar iPSCs a partir de células somáticas de tipo mesenquimal, se abordarán de manera resumida los procesos que ocurren a lo largo de las tres principales trayectorias celulares permisivas, que incluyen: la iniciación, la maduración y la estabilización (**Figura I4**).

La fase de iniciación a partir de fibroblastos se caracteriza principalmente por la extinción del programa transcripcional somático y la activación del epitelial, proceso conocido como transición mesenquimal-epitelial o MET (Hansson et al., 2012; Li et al., 2010; Pei et al., 2019; Samavarchi-Tehrani et al., 2010) (**Figura I4**). Este proceso va acompañado de cambios morfológicos y metabólicos en la población celular, así como de un aumento en su proliferación (Buganim et al., 2013; Mikkelsen et al., 2008; Smith et al., 2010; Smith et al., 2016; Stadtfeld et al., 2008). Posteriormente, la fase de maduración se caracteriza por una serie de cambios moleculares que promueven la activación de la expresión de ciertos genes que participan en la regulación del estado celular pluripotente (David y Polo, 2014; Polo et al., 2012; Samavarchi-Tehrani et al., 2010) (**Figura I4**). A diferencia de la MET, esta secuencia de cambios moleculares no es experimentada por todas las células sometidas al proceso de SCR, ya que la mayoría de ellas se vuelven refractarias a la reprogramación una vez iniciado el proceso (David y Polo, 2014; Polo et al., 2012). Este periodo es más largo que la fase de iniciación y puede

dividirse a su vez en diferentes estados intermedios (David y Polo, 2014) (**Figura I4**). Finalmente, estos cambios moleculares propician la fase de estabilización en la que se produce la activación de la red central de pluripotencia, entre otros eventos moleculares (Buganim et al., 2013) (**Figura I4**).

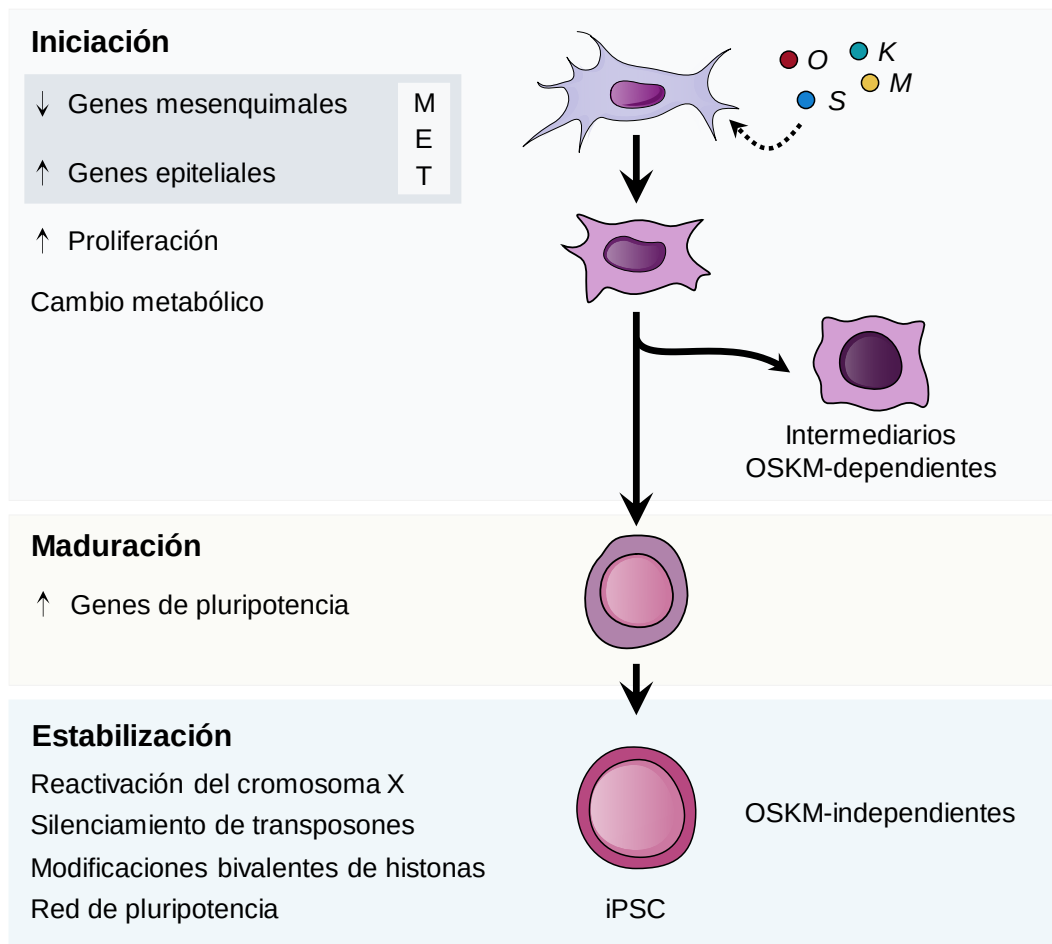


Figura I4. Fases permisivas del proceso de SCR mediado por OSKM a partir de células somáticas de origen mesenquimal. En la etapa de iniciación tiene lugar la MET. Esto implica la supresión del programa celular somático o mesenquimal y la activación del programa epitelial, acompañado de cambios en el metabolismo y proliferación celulares. Posteriormente, solo un grupo reducido de células logrará alcanzar la fase de maduración, caracterizada principalmente por la activación de la expresión de genes relacionados con el estado celular pluripotente. Finalmente, durante la fase de estabilización las células adquieren la capacidad de autorrenovarse de manera independiente de la expresión ectópica de OSKM. En esta última etapa, eventos críticos como la reactivación del cromosoma X, la represión de elementos transponibles y la deposición de modificaciones bivalentes de histonas, aseguran la consolidación de un estado pluripotente funcional. Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

3.2.2.1. Fase de iniciación

Después de la transducción con OSKM, tanto en fibroblastos humanos como de ratón, se produce un aumento de la proliferación celular y se inicia la MET (Buganim et al., 2013; Samavarchi-Tehrani et al., 2010; Takahashi et al., 2014) (**Figura I4**). Este proceso comienza con la inhibición del programa transcripcional somático, que incluye genes como *Snai1*, *Snai2*, *Twist1*, *Twist2*, *Zeb1* o *Zeb2* (Brambrink et al., 2008; Mikkelsen et al., 2008; Stadtfeld et al., 2008). Posteriormente, se activa el programa transcripcional epitelial mediante el aumento de la expresión de genes como *Cdh1* y *Epcam*, entre otros (David y Polo, 2014; Li

et al., 2010; Samavarchi-Tehrani et al., 2010; Takahashi y Yamanaka, 2015). Múltiples evidencias muestran que la MET es una etapa esencial durante la SCR, ya que su inhibición conlleva una reducción significativa en la eficiencia de reprogramación (Hoffding y Hyttel, 2015; Li et al., 2010; Pei et al., 2019). En consonancia con esto, varios estudios han documentado procesos similares a la MET durante la reprogramación de otros tipos celulares hacia iPSCs, como en linfocitos B (Di Stefano et al., 2014) o células precursoras hematopoyéticas (Gao et al., 2016; Pei et al., 2019). Estos hallazgos apoyan que la adquisición de la identidad epitelial es un requisito fundamental para el establecimiento de la pluripotencia a partir de células somáticas de origen mesenquimal (Apostolou y Stadtfeld, 2018; Deng et al., 2021; Pei et al., 2019).

3.2.2.2. Fases de maduración y estabilización

La transición de la fase de iniciación a la de maduración representa un punto limitante del proceso de SCR, ya que ocurre únicamente en el subconjunto de células reprogramables (Buganim et al., 2013; David y Polo, 2014; Tanabe et al., 2013). La fase de maduración coincide con una segunda serie de cambios moleculares relevantes, destacándose por el inicio de la expresión de genes asociados a la pluripotencia (David y Polo, 2014; Polo et al., 2012; Samavarchi-Tehrani et al., 2010) (**Figura I4**). Aunque la expresión de algunos marcadores pluripotentes como *Fbxo15*, *Sall4*, *Oct4* endógeno, *Nanog* y *Esrrb*, comienza durante la fase de maduración, otros como *Sox2* endógeno o *Dppa4*, se expresan posteriormente en la fase de estabilización (David y Polo, 2014; Samavarchi-Tehrani et al., 2010; Stadtfeld et al., 2008).

A diferencia de la extinción relativamente eficiente del programa somático, el establecimiento del programa pluripotente se considera un proceso variable y prolongado, con la posibilidad de ser dividido en estados intermedios (David y Polo, 2014). Sin embargo, con la iniciación de la expresión de *Nanog*, la variabilidad celular individual disminuye drásticamente, sugiriendo un modelo en el que una fase inicial "estocástica", que involucra la activación de algunos marcadores pluripotentes, es seguida por una fase "determinista" que conduce a la activación de la red reguladora de pluripotencia (Buganim et al., 2013).

Una característica esencial del proceso de SCR es la adquisición, por parte de las iPSCs, de la capacidad de autorrenovarse indefinidamente de manera independiente de la expresión ectópica de OSKM (David y Polo, 2014; Golipour et al., 2012). Durante la fase de estabilización se producen cambios que aseguran el mantenimiento de las iPSCs en ausencia de la expresión ectópica de factores de reprogramación, tras la adquisición del estado pluripotente (Brambrink et al., 2008; David y Polo, 2014; Maherali et al., 2007; Okita et al., 2007; Stadtfeld et al., 2008) (**Figura I4**). Diversas evidencias sugieren que el restablecimiento gradual de la red reguladora de pluripotencia se favorece tras la activación de un subconjunto de genes durante la fase de maduración y que requiere del silenciamiento simultáneo de la expresión ectópica de OSKM (Apostolou y Stadtfeld, 2018; Deng et al., 2021; Golipour et al., 2012).

Finalmente, eventos adicionales son necesarios para el establecimiento de una pluripotencia funcional, además de la activación independiente de OSKM de la red de pluripotencia. Estos eventos incluyen la desmetilación del ADN, la reactivación del cromosoma X en células femeninas (David y Polo, 2014; Maherali et al., 2007; Stadtfeld et al., 2008), la reactivación de la telomerasa (Stadtfeld et al., 2008) y el restablecimiento de las modificaciones bivalentes de histonas en genes del desarrollo (Smith et al., 2016; Wernig et al., 2007) (**Figura I4**), entre otros eventos moleculares.

3.3 MODIFICACIONES QUÍMICAS EN ARN

3.3.1 Visión general

Las modificaciones químicas del ARN, conocidas como modificaciones epitranscriptómicas, constituyen un mecanismo clave en la regulación espacial y temporal de programas de expresión génica (Boccaletto et al., 2018; Kumari et al., 2021). Actualmente, se conocen más de 170 modificaciones químicas presentes en diversas especies de ARN, incluyendo el ARN transferente (ARNt), el ARN ribosómico (ARNr), el ARN mensajero (ARNm), así como moléculas de ARN no codificantes (ARNnc) (Boccaletto et al., 2018; Boo y Kim, 2020) (**Figura I5**).

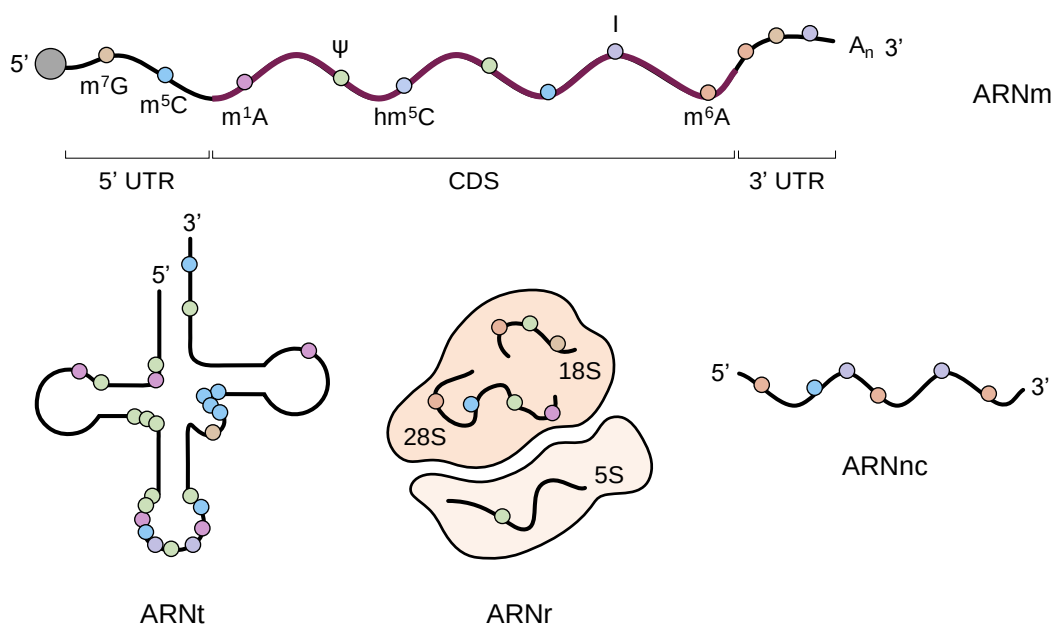


Figura I5. Las modificaciones epitranscriptómicas ocurren en diversas especies de ARN. Representación esquemática de las modificaciones de ARN más estudiadas en las moléculas de ARNm, ARNt, ARNr y ARNnc. Para el ARNm, se destacan sus regiones principales: región 5' UTR, región codificante (CDS) y región 3' UTR. En el ARNr, se indican los tipos de ARN que constituyen cada una de las subunidades del ribosoma: los ARNr 28S y 18S forman parte de la subunidad mayor y el 5S forma parte de la subunidad menor. Se muestran las siguientes modificaciones epitranscriptómicas: N7-metilguanósina (m^7G), N1-metiladenósina (m^1A), 5-metilcitosina (m^5C), 5-hidroximetilcitosina (hm^5C), Inosina (I) y pseudouridina (ψ). Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

A pesar de que la existencia de ribonucleótidos modificados se documentó hace más de 70 años (Cohn y Volkin, 1951), las funciones de la mayoría de estas modificaciones aún permanecen, en gran medida, desconocidas. La mayor parte de estas alteraciones químicas han sido identificadas en moléculas de ARN muy abundantes, como el ARNr o el ARNt, siendo este último la especie de ARN más enriquecida en residuos modificados (Roundtree et al., 2017) (**Figura I5**). Sin embargo, los avances recientes en tecnologías de secuenciación de alto rendimiento y otras aproximaciones, como la espectrometría de masas, han permitido la detección y caracterización de residuos modificados en especies de ARN menos abundantes, como el ARNm, entre otras (Kadumuri y Janga, 2018; Lewis et al., 2017; Ofusa et al., 2022; Trixl y Lusser, 2019).

La epitranscriptómica desempeña un papel crítico en la regulación del metabolismo del ARN. Estas modificaciones pueden influir directamente en su estructura favoreciendo o perturbando determinadas interacciones intramoleculares, así como en su interacción con otras

moléculas, especialmente con proteínas (Boccaletto et al., 2018). De esta manera, las modificaciones químicas del ARN impactan en diversos procesos moleculares, como la transcripción, la traducción del ARNm, el corte y empalme ("*splicing*"), la exportación del ARN, su localización y su degradación (Boo y Kim, 2020; Roundtree et al., 2017; Shi et al., 2020). Estas alteraciones en el metabolismo del ARN se traducen en cambios en la regulación de programas de expresión génica que participan en una amplia variedad de procesos biológicos, tales como el establecimiento de la identidad celular, la diferenciación o la homeostasis, entre otros (Li y Mason, 2014; Roundtree et al., 2017).

Generalmente, tal y como se ha descrito para una de las modificaciones químicas del ARN más estudiadas, la N6-metiladenosina (m^6A), el destino de los transcritos modificados está determinado por la acción coordinada de tres grupos de proteínas: las proteínas modificadoras de ARN o "*writers*", responsables de la deposición de la marca epitranscriptómica en las moléculas de ARN; las proteínas de unión a ARN o "*readers*", encargadas de su reconocimiento; y las proteínas borradoras o "*erasers*", que eliminan los residuos modificados (Boo y Kim, 2020; Kadumuri y Janga, 2018; Kumari et al., 2021). Mientras que algunas de estas proteínas actúan de manera independiente, otras lo hacen mediante la formación de complejos compuestos por subunidades accesorias (Liu et al., 2014). La regulación dinámica de las modificaciones de ARN, orquestada por los "*writers*", "*readers*" y "*erasers*", garantiza, como en el caso de m^6A , su reversibilidad, permitiendo una rápida respuesta postranscripcional a distintos estímulos o variaciones ambientales. De hecho, la exposición a determinados agentes exógenos y endógenos también puede provocar modificaciones del ARN sin la participación de otras proteínas (Boccaletto et al., 2022; Boo y Kim, 2020; Roundtree et al., 2017; Wang y He, 2014).

Tanto la mutación de enzimas modificadoras de ARN, como la alteración de los patrones de modificación de esta molécula, se encuentran vinculadas a diversas enfermedades humanas, incluyendo el cáncer, las enfermedades cardiovasculares y los trastornos metabólicos e inmunitarios, entre otros (Jonkhout et al., 2017; Kadumuri y Janga, 2018). Por consiguiente, se requiere un conocimiento más profundo de las propiedades de los transcritos modificados y de las enzimas responsables de su modificación con el fin de diseñar estrategias terapéuticas novedosas y específicas contra enfermedades que, hasta el momento, carecen de cura.

3.3.2 Distribución, funciones y regulación de la marca m^5C

La 5-metilcitosina (referida como 5mC en ADN y m^5C en ARN), que implica la adición de un grupo metilo (CH_3) al carbono en la posición número 5 de la base nitrogenada citosina (**Figura I6**), es una modificación ampliamente conocida por su papel central como marca epigenética en el ADN. Sus funciones y mecanismos de regulación en este contexto han sido estudiados de manera exhaustiva durante décadas. Sin embargo, a pesar de su extensa distribución en diversos organismos, la relevancia fisiológica y patológica de la marca m^5C en el ARN, que será objeto de estudio en esta tesis doctoral, continúa sin ser completamente esclarecida (Gao y Fang, 2021; Garcia-Vilchez et al., 2019; Nombela et al., 2021; Shi et al., 2020; Trixl y Lusser, 2019).

Los avances recientes en tecnologías como la secuenciación de tercera generación, la espectrometría de masas, la inmunoprecipitación o la secuenciación por bisulfito han permitido la detección de m^5C en diversas especies de ARN (Amort et al., 2017; Hussain et al., 2013a; Lucas et al., 2023; Parker et al., 2020; Su et al., 2014; Trixl et al., 2019) (**Figura I5**). La presencia de m^5C ha sido observada en ARNr, ARNt, ARNm, y ARNnc como el microARN

(miARN) o el ARN derivado de potenciadores o *enhancers* (ARNe), entre otros (Gao y Fang, 2021; Trixl y Lusser, 2019) (**Figura I5**).

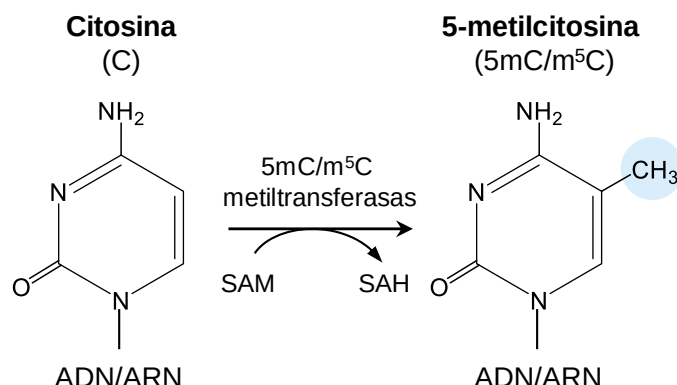


Figura I6. La modificación 5-metilcitosina se encuentra presente en moléculas de ADN y ARN. Representación simplificada del proceso de metilación del carbono situado en la posición 5 de la base nitrogenada citosina en ADN y ARN. El proceso es catalizado por grupos de enzimas (5mC/m⁵C-metiltransferasas) que emplean el cosustrato S-adenosilmetionina (SAM) como donador de grupos metilo, provocando así su desmetilación y dando lugar a la molécula S-adenosil L-homocisteína (SAH). Se señala con un círculo azul el grupo metilo (CH₃) en la posición 5 de la citosina. Esta imagen fue creada con el programa ChemDraw (v22.2).

Sin embargo, la localización y abundancia de la marca m⁵C varían entre diferentes tipos de ARN y organismos. Mientras que el ARNt está muy enriquecido en m⁵C en organismos eucariotas, esta modificación es menos abundante en ARNt procariota (Amort et al., 2017; Legrand et al., 2017; Motorin et al., 2010; Squires et al., 2012; Yang et al., 2017). Por otro lado, la presencia de m⁵C en el ARNm eucariota ha sido objeto de controversia debido a las diferencias en los protocolos de detección y análisis. Aunque varios estudios confirman la presencia de m⁵C en ARNm, las discrepancias en los métodos han dificultado la identificación precisa de los sitios de m⁵C, incluso después del análisis de los mismos tipos celulares (Amort et al., 2017; Huang et al., 2019; Legrand et al., 2017; Squires et al., 2012; Yang et al., 2017). Es importante destacar que la distribución de m⁵C también varía entre células de organismos eucariotas. Por ejemplo, en células de mamíferos, esta modificación química se localiza en la región codificante (CDS, por sus siglas en inglés) de las moléculas de ARNm, mostrando un enriquecimiento alrededor del sitio de inicio de la traducción (Amort et al., 2017; Yang et al., 2017), mientras que en embriones de pez cebra los sitios con m⁵C se distribuyen de manera homogénea a lo largo de la CDS (Yang et al., 2019c).

La modificación m⁵C desempeña funciones críticas en varios aspectos del metabolismo del ARN, incluyendo su estabilidad (Chen et al., 2019; Schaefer et al., 2010; Tuorto et al., 2012; Yang et al., 2019c; Zhang et al., 2018), transporte (Yang et al., 2019a; Yang et al., 2017) y traducción (Tuorto et al., 2015; Tuorto et al., 2012) (**Figura I7**). Además, es fundamental para la regulación de procesos biológicos como la respuesta funcional al estrés (Blanco et al., 2014; Durdevic et al., 2013b; Lewinska et al., 2018; Schaefer et al., 2010), el desarrollo embrionario (Rai et al., 2007; Yang et al., 2019c), la diferenciación (Hussain et al., 2013b; Tuorto et al., 2015; Tuorto et al., 2012), y la herencia de fenotipos entre generaciones (Kiani et al., 2013; Zhang et al., 2018), entre otros (**Figura I7**). De esta manera, la alteración de la función de las enzimas que regulan el metabolismo de la marca m⁵C se asocia con procesos patológicos, entre los que se incluyen trastornos neurológicos y cáncer (Blanco et al., 2016; Chen et al., 2019; Frye et al., 2010; Khan et al., 2012).

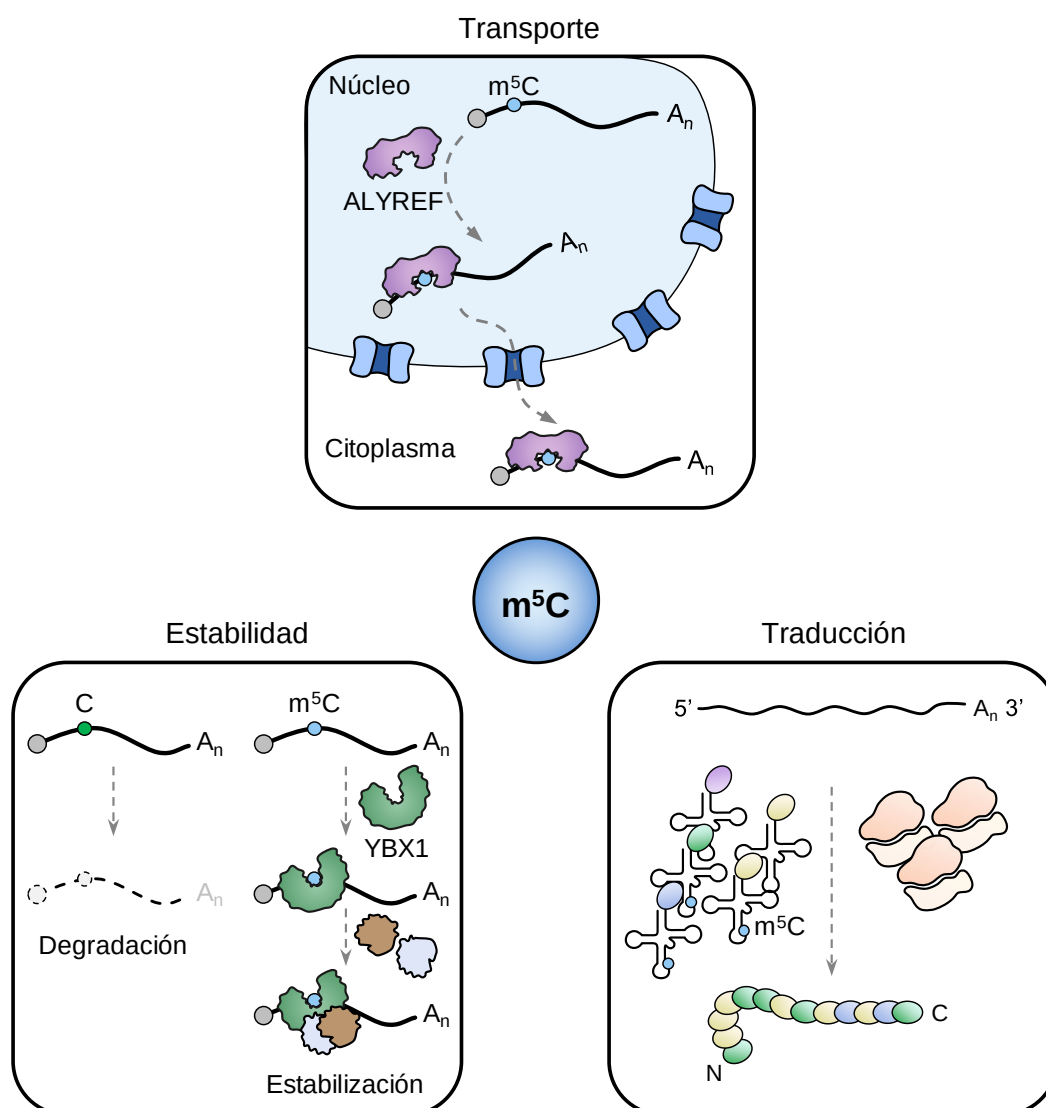


Figura I7. La modificación m^5C interviene en el metabolismo del ARN. Representación esquemática de las principales funciones de la modificación m^5C en la célula. La presencia de esta marca ha sido correlacionada positivamente con una mayor estabilidad de las moléculas de ARN, ya que permite el reconocimiento de estas moléculas por proteínas (por ejemplo, YBX1) que dificultan su degradación. Además, también se ha observado que su presencia favorece el transporte mediado por proteínas (por ejemplo, ALYREF) de moléculas de ARNm del núcleo al citoplasma. Por otro lado, la modificación m^5C también regula la traducción promoviendo la estabilidad de los ARNt, así como la distinción entre codones muy similares y por lo tanto, favoreciendo la fidelidad del proceso. Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

Hasta el momento, se han identificado dos grupos de m^5C -metiltransferasas de ARN, altamente conservadas a nivel evolutivo, que catalizan la deposición de m^5C utilizando el cofactor S-adenosilmetionina (SAM) como donador de grupos metilo (Song et al., 2022). Estos grupos incluyen la familia de proteínas NOL1/NOP2/sun (NSUN), compuesta por 7 miembros en mamíferos, desde NSUN1 hasta NSUN7; y la proteína DNMT2, objeto principal de estudio en esta tesis doctoral, también conocida como TRDMT1, que pertenece a la familia de las ADN metiltransferasas (DNMTs) (**Figura I8**).

Las proteínas NSUN han sido objeto de estudio en diversos contextos celulares, lo que ha permitido identificar su distinta especificidad de sustratos (Trixl y Lusser, 2019) (**Tabla 1**). Por ejemplo, las proteínas NSUN1 y NSUN5, esencialmente localizadas en el núcleo, han sido

identificadas principalmente como m^5C -metiltransferasas exclusivas de ARNr (Heissenberger et al., 2019; Heissenberger et al., 2020; Liao et al., 2022). Por otro lado, NSUN3 y NSUN4, actúan sobre ARNt mitocondrial (ARNt-mt) (Delaunay et al., 2022; Haag et al., 2016; Nakano et al., 2016) y ARNr mitocondrial (ARNr-mt) (Metodieiev et al., 2014), respectivamente. NSUN2 y NSUN6 participan en la metilación de ARNt citoplasmáticos, y mitocondriales en el caso de NSUN2 (Van Haute et al., 2019), junto con DNMT2, pero con diferencias en la especificidad de moléculas y residuos a modificar (Selmi et al., 2021; Tuorto et al., 2012). Además, también catalizan la deposición de m^5C en otros tipos de ARN como ARNm y ARNnc (Chen et al., 2019; Huang et al., 2019; Selmi et al., 2021), mientras que NSUN7 lo hace en ARNe (Aguilo et al., 2016) y ARNm (Ortiz-Barahona et al., 2023).

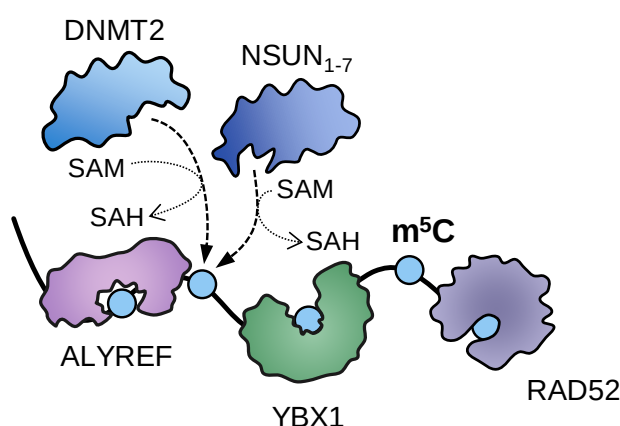


Figura I8. La deposición de m^5C en ARN es catalizada por DNMT2 y la familia de proteínas NSUN. Esquema simplificado en el que se señalan las principales proteínas reguladoras del metabolismo de la marca m^5C . Tanto DNMT2 como la familia de proteínas NSUN se encargan de su catálisis en distintos tipos de ARN, empleando la molécula SAM como donadora de grupos metilo. Por otro lado, se muestran tres "readers" de m^5C en mamíferos: ALYREF, YBX1 y RAD52. Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

A través de la modificación de distintos tipos de ARN, las proteínas NSUN ejercen una influencia directa en procesos celulares esenciales como la transcripción y la traducción (Delaunay et al., 2023; Garcia-Vilchez et al., 2019; Nombela et al., 2021). La alteración de la función de estas proteínas se ha asociado con defectos graves en el desarrollo y la maduración de órganos, así como con diversas enfermedades (Delaunay y Frye, 2019; Delaunay et al., 2023; Frye et al., 2018). La eliminación de NSUN3 o NSUN4, por ejemplo, conduce a eventos letales durante el desarrollo embrionario en ratones (Metodieiev et al., 2014; Murakami et al., 2023). Asimismo, la pérdida de la actividad catalítica de NSUN3 también está vinculada a defectos en la capacidad de diferenciación de células pluripotentes (Trixl et al., 2018). En cuanto a NSUN2, su pérdida de función se asocia comúnmente con eventos de neurodegeneración y discapacidad intelectual (Blaze et al., 2021; Khan et al., 2012). Además, diversos estudios han documentado la relación entre cambios en la función de las proteínas NSUN y eventos de transformación oncogénica (Chen et al., 2019; Delaunay et al., 2022; Frye et al., 2010; Li et al., 2023; Nombela et al., 2021; Ortiz-Barahona et al., 2023; Selmi et al., 2021).

Contrariamente, los sustratos y procesos celulares afectados por la enzima DNMT2 no han sido estudiados de manera tan extensa. Esto posiblemente se debe a que su pérdida de función no ha sido asociada hasta ahora con fenotipos graves (Goll et al., 2006; Kiani et al., 2013; Tuorto et al., 2015). Sin embargo, esta proteína plantea una interesante incógnita, ya que es la DNMT más conservada a nivel evolutivo, encontrándose presente tanto en procariotas unicelulares (Fisher et al., 2006) como en organismos eucariotas (Lyko, 2018). Además, es la

única DNMT capaz de actuar sobre ARN en mamíferos (Lyko, 2018). De este modo, la alta conservación filogenética de DNMT2 sugiere un papel crítico en procesos biológicos fundamentales que aún no se han explorado de manera exhaustiva.

Proteína	Sustrato	Localización celular
DNMT2	ARNt (C38) ARNnc ARNm	Citoplasma/Núcleo
NSUN1	ARNr (28S)	Citoplasma/Núcleo
NSUN2	ARNt (C34, C48-C50) ARNm ARNnc	Citoplasma/Núcleo/Mitocondria
NSUN3	ARNt-mt (C34)	Mitocondria
NSUN4	ARNr-mt (12S)	Mitocondria
NSUN5	ARNr (28S)	Núcleo
NSUN6	ARNt (C72) ARNm ARNnc	Citoplasma/Núcleo
NSUN7	ARNe ARNm	Núcleo

Tabla 1. DNMT2 y la familia de proteínas NSUN catalizan la deposición de m⁵C en diferentes sustratos de ARN. Tabla resumen de los principales tipos de ARN modificados por DNMT2 y las proteínas NSUN, junto con la localización celular de cada enzima. Se incluyen también las posiciones específicas de los ARNt donde cada proteína cataliza la modificación m⁵C.

Otras proteínas que desempeñan un papel crucial en la regulación de la marca m⁵C son aquellas encargadas de su reconocimiento y eliminación. Hasta la fecha, algunas enzimas de reconocimiento de m⁵C incluyen: YBX1, que participa en la estabilización de transcritos marcados con m⁵C (Chen et al., 2019); ALYREF, que reconoce transcritos marcados con m⁵C y facilita su exportación desde el núcleo al citoplasma celular (Yang et al., 2017); y RAD52, que identifica m⁵C en híbridos de ARN y ADN formados después de eventos de rotura de doble hebra, promoviendo la recombinación homóloga para su reparación (Chen et al., 2020) (**Figura 18**).

En cuanto a los factores responsables de la eliminación activa de la modificación m⁵C, el conocimiento aún es limitado. Hasta el momento, se ha demostrado que m⁵C puede someterse a un proceso oxidativo, dando lugar a 5-hidroximetilcitosina (hm⁵C) y 5-formilcitosina (f⁵C) en ARN (Fu et al., 2014; Huang et al., 2016; Huber et al., 2015; Zhang et al., 2016). Este proceso de eliminación de m⁵C en ARN se ha descrito como mediado por la acción directa de miembros de la familia *Ten-Eleven Translocation* (TET) (Delatte et al., 2016; Fu et al., 2014; Lan et al., 2020), en particular TET2, que da lugar a hm⁵C (Garcia-Outeiral et al., 2020; Guallar et al., 2018; He et al., 2021a). Otra enzima implicada es ALKBH1, responsable de hm⁵C y f⁵C (Arguello et al., 2022).

3.4 DNMT2

3.4.1 Filogenia, estructura y actividad catalítica

La proteína DNMT2 destaca como un miembro atípico de la familia de las DNMTs, que en mamíferos se compone de cinco proteínas: DNMT1, DNMT2, DNMT3A, DNMT3B y DNMT3L (Lyko, 2018) (**Figura I9**). A pesar de haber sido inicialmente clasificada como parte de esta familia debido a su elevada homología en términos de secuencia y estructura con el dominio catalítico de las DNMT (Dong et al., 2001; Jurkowski y Jeltsch, 2011) (**Figura I9**), esta proteína difiere significativamente de las demás enzimas de este grupo.

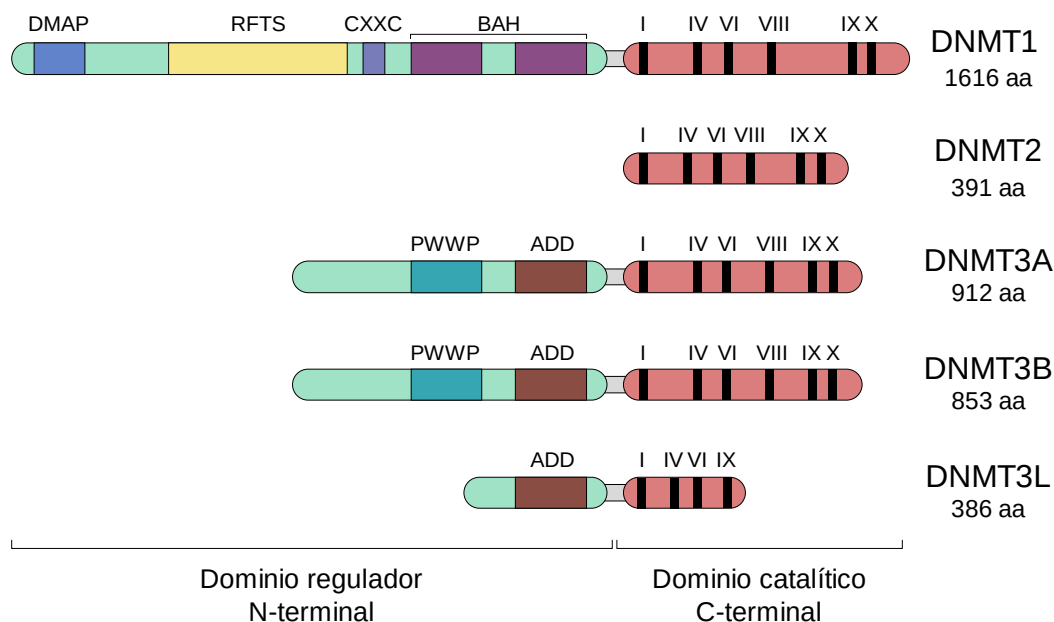


Figura I9. La familia de proteínas DNMT. Esquema simplificado de la familia de proteínas DNMT, compuesta por cinco miembros en humanos: DNMT1, DNMT2, DNMT3A, DNMT3B y DNMT3L. En la figura, se destacan el dominio regulador N-terminal (verde), crucial para la unión al ADN y ausente en DNMT2; y el dominio catalítico C-terminal (rojo), implicado en el proceso de metilación. También se muestran algunos subdominios presentes en estas proteínas. En DNMT1, el motivo N-terminal contiene varios subdominios conservados que facilitan interacciones moleculares: DMAP1, crucial para mediar interacciones proteína-proteína; RFTS y BAH, necesarios para dirigir a DNMT1 a los focos de replicación; y CXXC, que media la unión a ADN no metilado. Los subdominios PWWP y ADD, presentes en las proteínas DNMT3, median su interacción con otras proteínas, como histonas. DNMT3L es una variante catalíticamente inactiva de DNMT3, que carece de la parte N-terminal del dominio regulador (incluido el subdominio PWWP) y de la parte C-terminal del dominio catalítico. Los motivos catalíticos (en negro) están muy conservados en todas las DNMTs, excepto en DNMT3L. También se indica el número de aminoácidos de cada proteína. Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

A diferencia de las otras proteínas de la familia DNMT, DNMT2 carece del dominio regulador y de unión al ADN, localizado en el extremo N-terminal de estas proteínas, así como de los motivos esenciales para llevar a cabo su metilación (Lyko, 2018). De este modo, su estructura se compone exclusivamente de un dominio catalítico (Lyko, 2018) (**Figura I9**). Aunque ensayos iniciales realizados en ESCs de ratón indicaron la ausencia de cambios en los patrones de metilación del ADN en células mutantes para DNMT2 (Okano et al., 1998), investigaciones posteriores han revelado niveles residuales de metilación del ADN por parte de esta enzima en ensayos *in vitro* con proteína purificada (Hermann et al., 2003) o en oligonucleótidos de ADN de cadena simple, también *in vitro*, mediante un sistema compuesto de ARN guía (Kaiser et al., 2017). Además, se propuso que la participación de la actividad

catalítica de DNMT2 sobre el ADN está involucrada en el control de la expresión de retrotransposones endógenos en células somáticas de *Drosophila melanogaster* durante la embriogénesis temprana (Phalke et al., 2009), aunque este estudio ha sido puesto en duda por falta de evidencias sólidas (Schaefer y Lyko, 2010). En consecuencia, a pesar de que DNMT2 cataliza eficientemente la metilación del ARN, como se detallará a continuación, la actividad metiltransferasa de ADN de esta enzima en ciertos organismos que carecen de las otras enzimas DNMT (por ejemplo, *Drosophila melanogaster*) sigue siendo motivo de controversia en la actualidad.

Fue en 2006 cuando el grupo del Dr. Timothy H. Bestor demostró que la actividad catalítica principal de DNMT2 radica en la metilación de moléculas de ARNt (Goll et al., 2006). Este estudio postuló por primera vez el papel de DNMT2 como una m⁵C-metiltransferasa de ARN en organismos como *Arabidopsis thaliana*, *Drosophila melanogaster* y *Mus musculus*. En estos se demostró que su sustrato principal era la citosina en la posición 38 (C38), situada en el bucle del anticodón, de las moléculas de ARNt que portaban ácido aspártico (ARNt-Asp) (Goll et al., 2006) (**Figura I10**). Posteriormente, estudios *in vivo* en los mismos organismos (Schaefer et al., 2010; Tuorto et al., 2012), así como en células humanas (Khoddami y Cairns, 2013), revelaron que DNMT2 también cataliza la metilación de la posición C38 en otros ARNt, como ARNt-Gly, portador del aminoácido glicina; y ARNt-Val, portador del aminoácido valina. Además, en procariotas y eucariotas unicelulares, se ha observado la actividad de DNMT2 en algunos ARNt que llevan ácido glutámico, como ARNt-Glu (Shanmugam et al., 2014). Asimismo, trabajos recientes *in vitro* señalan la capacidad de DNMT2 para llevar a cabo la metilación de otras moléculas de ARNt, como ARNt-Ala (Li et al., 2021).

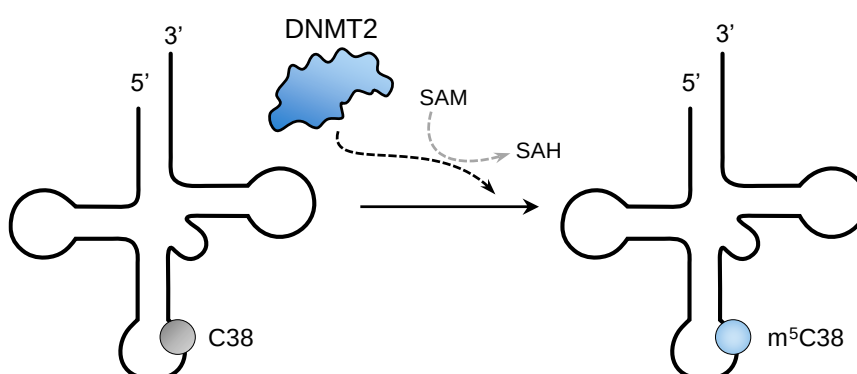


Figura I10. DNMT2 cataliza la deposición de m⁵C en la citosina 38 de algunos ARNt.

Representación simplificada de la actividad catalítica de la enzima DNMT2 sobre la posición C38, próxima al anticodón, de las moléculas de ARNt. Esta proteína cataliza la deposición de m⁵C empleando el cosustrato SAM como donador de grupos metilo. Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

Cabe destacar que, aunque la actividad principal de DNMT2 radica en la deposición de la marca m⁵C en ARNt, varios estudios han evidenciado la presencia de m⁵C mediada por DNMT2 en otros ARN, como ARNm (Chen et al., 2020; Dev et al., 2017; Khoddami y Cairns, 2013), o ARNnc (Ghanbarian et al., 2016; Khoddami y Cairns, 2013), siendo crucial para la estabilización de algunas de estas moléculas.

Como se mencionó anteriormente, DNMT2 muestra una mayor relación estructural con las proteínas DNMT que con las otras m⁵C-metiltransferasas de ARN de la familia NSUN (Jurkowski y Jeltsch, 2011) (**Figura I11a**). Aunque DNMT2 es la única DNMT capaz de actuar sobre ARN, emplea para ello un mecanismo catalítico muy similar al de las proteínas de esta

familia, diferente al utilizado por las proteínas NSUN (Jurkowski et al., 2008; Liu y Santi, 2000).

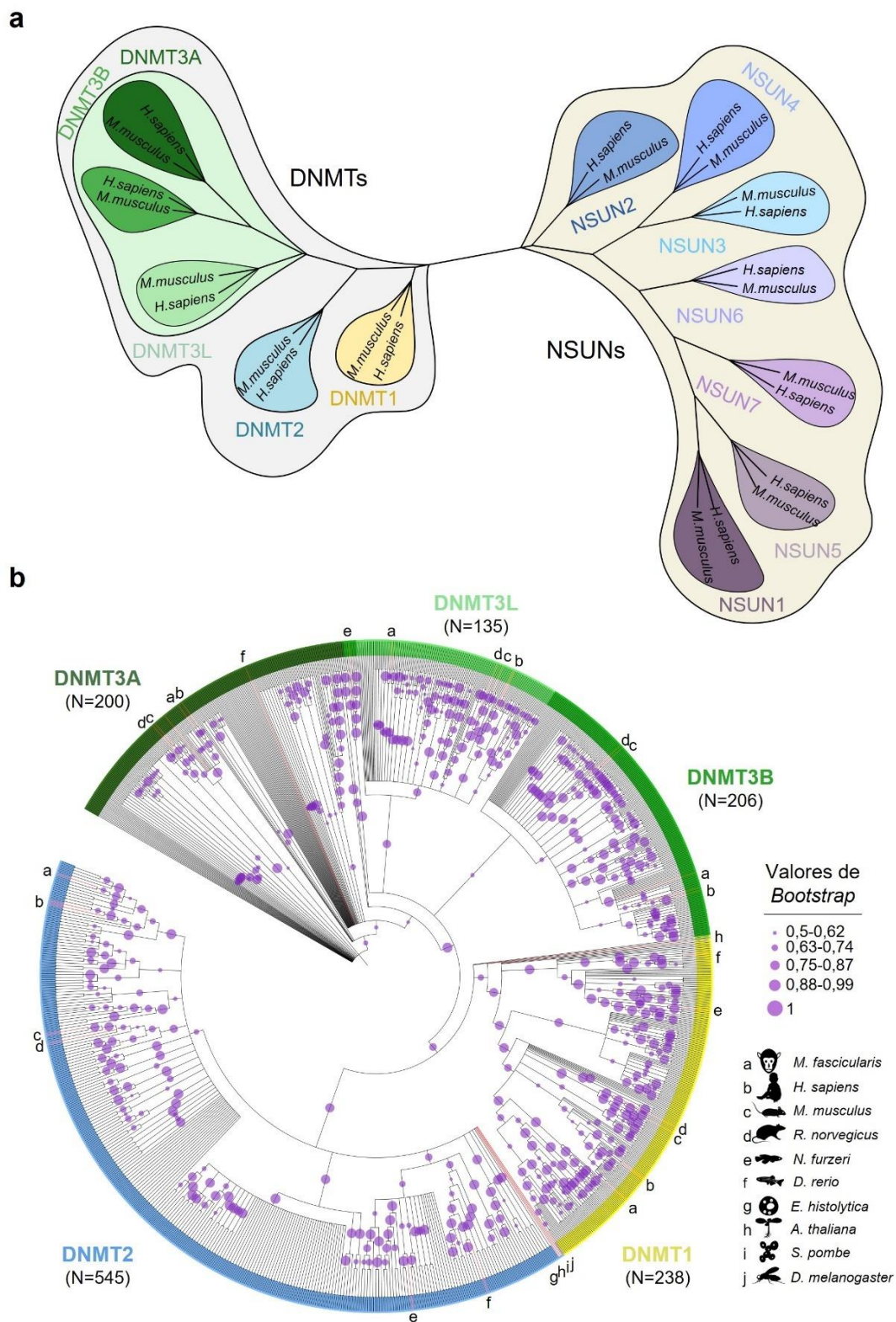


Figura I11. Análisis de la divergencia y de la relación evolutiva de la proteína DNMT2. a) Esquema que muestra las relaciones filogenéticas entre los miembros de las familias de proteínas DNMT y NSUN en *Homo sapiens* y *Mus*

musculus, en base a su secuencia proteica. b) Árbol filogenético consenso construido a partir de 1000 simulaciones (*bootstrap*) de este análisis, en el que se muestra el grado de conservación de las proteínas DNMT en múltiples especies de vertebrados. El valor de *bootstrap* hace referencia al porcentaje de coincidencia de cada una de las ramas del árbol respecto a las 1000 pruebas. También se muestra el número de organismos que presentan cada una de estas proteínas (N). Las letras en minúscula señalan aquellos organismos habitualmente empleados como modelo de investigación (indicados a la derecha). Los análisis filogenéticos de los paneles a y b fueron elaborados mediante Clustal Omega y MEGA X, respectivamente. La imagen del panel a fue creada en Power Point y la del panel b mediante la herramienta web iTOL (Letunic y Bork, 2021). Para más detalles, consultar la sección 5.29 de materiales y métodos.

Por otro lado, DNMT2 es la DNMT mejor conservada filogenéticamente, ya que está presente en diversos organismos, desde procariotas y eucariotas unicelulares (por ejemplo, *Entamoeba histolytica* y *Schizosaccharomyces pombe*, respectivamente), hasta organismos pluricelulares (Jurkowski y Jeltsch, 2011) (**Figura I11b**). Además, varios organismos, comúnmente conocidos como organismos "*Dnmt2-only*", conservan a DNMT2 como el único miembro de esta familia (por ejemplo, *Entamoeba histolytica*, *Schizosaccharomyces pombe* y *Drosophila melanogaster*, entre otros) (Jeltsch et al., 2017), lo que destaca su alta relevancia a nivel evolutivo.

3.4.2 Funciones biológicas de DNMT2

En las últimas décadas, varios estudios han confirmado que la metilación de C38 mediada por DNMT2 desempeña un papel crítico en diversos procesos celulares, entre los que se encuentran el procesamiento de ARNt, la traducción, la diferenciación celular y la herencia transgeneracional.

3.4.2.1 Procesamiento del ARNt

La metilación de la posición C38, catalizada por DNMT2 en distintos ARNt, influye en la estructura y estabilidad de estas moléculas. Esta modificación protege a los ARNt de su escisión y la consecuente generación de fragmentos derivados de ARNt (tsRNAs, por sus siglas en inglés), mediada por la endorribonucleasa angiogenina (ANG), que muestra una mayor afinidad por los residuos que carecen de m⁵C (Blanco et al., 2014) (**Figura I12**). La fragmentación de estos ARN resulta en la exposición de sitios de unión a proteínas, previamente ocultos en su estructura plegada, lo que facilita el establecimiento de nuevas interacciones (Schimmel, 2018). Específicamente, la escisión del ARNt en la región del anticodón genera fragmentos que participan en la formación de gránulos de respuesta a estrés y son capaces de inhibir la traducción en este contexto (Emara et al., 2010; Ivanov et al., 2011; Schimmel, 2018) (**Figura I12**).

La actividad endonucleasa de la enzima ANG, y en consecuencia la generación de tsRNAs, aumentan bajo condiciones de estrés celular (por ejemplo, estrés térmico o estrés oxidativo) (Durdevic et al., 2013b; Schaefer et al., 2010; Tuorto et al., 2012). De manera similar, la eliminación de DNMT2, y por ende de m⁵C38, también resulta en una mayor producción de estos fragmentos, incrementada frente a distintos tipos de estrés en diversos organismos (Genencher et al., 2018; Zhang et al., 2018). En este contexto, DNMT2 se desplaza del núcleo al citoplasma, donde forma parte de complejos de procesamiento del ARN y participa en el ensamblaje de gránulos de respuesta a estrés (Dev et al., 2017; Schaefer et al., 2010). También, en algunos organismos como *Drosophila melanogaster* o *Entamoeba histolytica*, se ha mostrado que DNMT2 regula la expresión de proteínas de choque térmico (HSP, por sus siglas en inglés), fundamentales en la protección frente a distintos tipos de estrés celular (Fisher et al., 2006; Lin et al., 2005; Thiagarajan et al., 2011). Además, varios estudios

han documentado cómo la ausencia de esta proteína compromete la esperanza de vida de *Drosophila melanogaster* bajo condiciones de estrés, mientras que su sobreexpresión la prolonga (Lin et al., 2005; Schaefer et al., 2010).

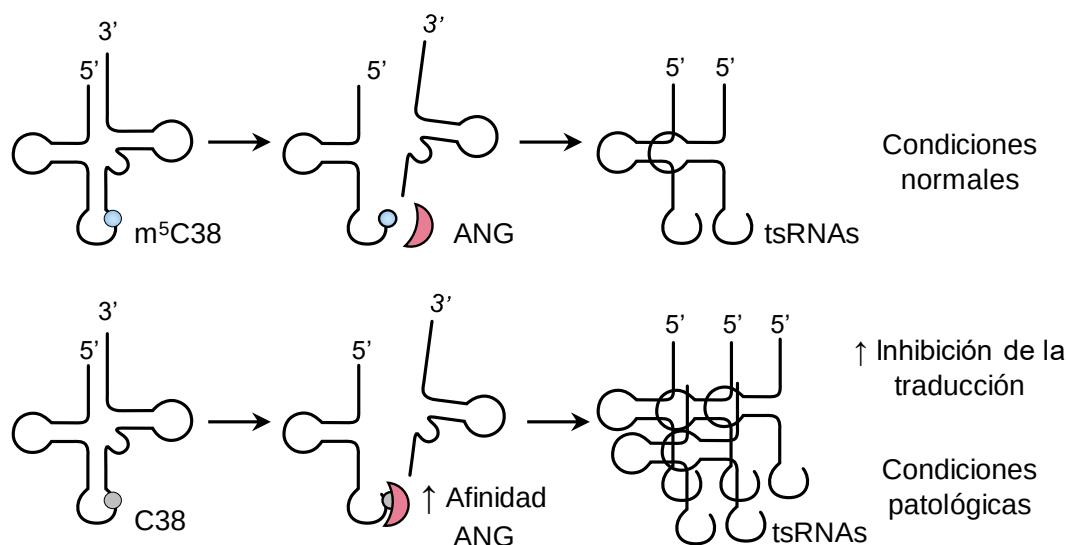


Figura I12. La ausencia de m⁵C38 promueve la generación de tsRNAs. La pérdida de la marca m⁵C en C38 (m⁵C38) conduce a un aumento en la generación de tsRNAs mediada por angiogenina (ANG), ya que la afinidad de esta enzima por nucleótidos no modificados se ve incrementada. La acumulación resultante de tsRNAs contribuye al aumento de eventos de inhibición de la traducción, además de favorecer la aparición de patologías asociadas (por ejemplo, la formación de tumores). Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

Por otro lado, tanto en *Drosophila melanogaster* (Durdevic et al., 2013b) como en humanos (Haussecker et al., 2010), se observó que los tsRNAs pueden modular la eficiencia de las rutas de silenciamiento mediadas por ARN pequeños de interferencia (siRNA, por sus siglas en inglés) a través de la interacción con proteínas como Dicer y Argonauta (Durdevic et al., 2013b; Haussecker et al., 2010). Por ejemplo, la eliminación de DNMT2 en *D. melanogaster*, bajo condiciones de estrés térmico, conduce a un aumento en la generación de tsRNAs capaces de inhibir la actividad de Dicer (Dcr-2) sobre ARN de doble cadena (Durdevic et al., 2013b). Esto provoca su acumulación y, en consecuencia, defectos en la regulación de genes dependientes de la vía de siRNA, como algunos implicados en la respuesta inmune (Durdevic et al., 2013a; Durdevic et al., 2013b). De hecho, la ausencia de DNMT2 se ha relacionado con una menor eficiencia en la respuesta inmune innata de *Drosophila melanogaster*, dependiente de esta vía (Durdevic et al., 2013a; Wang et al., 2006).

También se ha demostrado que los tsRNAs pueden intervenir en la regulación de procesos patológicos, como los asociados al cáncer. Un ejemplo es el papel antitumoral que presentan algunos tsRNAs en el contexto del cáncer de mama, ya que son capaces de capturar a la proteína YBX1 ("reader" de m⁵C), evitando así la estabilización de algunos oncogenes (Goodarzi et al., 2015). Sin embargo, estas moléculas también pueden favorecer la progresión del tumor en otros contextos. Este es el caso del fragmento tRF-Leu-CAG-3', capaz de potenciar la traducción de los ARNm codificantes de ribosomas *RPS28* y *RPS15* y cuya inhibición conlleva una reducción del tamaño de tumores hepatocelulares (Kim et al., 2017). Teniendo esto en cuenta y considerando que la expresión y actividad de DNMT2 varían significativamente en distintos tipos de tumor respecto al tejido no tumoral (Elhardt et al., 2015;

Schaefer et al., 2009; Sun et al., 2021), es posible que la actividad catalítica de esta proteína sea relevante para el avance o regresión de algunos tipos de cáncer (Elhardt et al., 2015).

3.4.2.2 Traducción y diferenciación celular

La capacidad de DNMT2 para proteger el ARNt de su fragmentación mediante la metilación de C38 influye en el proceso de traducción (Shanmugam et al., 2015; Tuorto et al., 2015) (**Figura I13**). Sin embargo, los niveles globales de síntesis de proteínas no se ven afectados en ausencia de DNMT2 o m⁵C38 (Tuorto et al., 2015; Tuorto et al., 2012). No obstante, la pérdida de función conjunta de DNMT2 y NSUN2 sí provoca una reducción significativa de la tasa global de traducción (Tuorto et al., 2012), sugiriendo un papel compensatorio entre ambas proteínas.

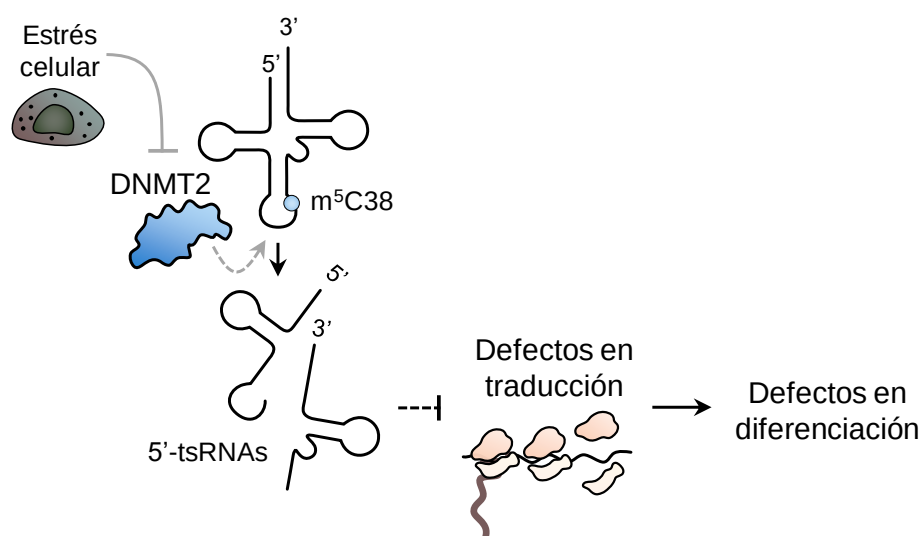


Figura I13. La función de DNMT2 es crítica para la correcta síntesis de proteínas. Los ARNt metilados en C38 desempeñan un papel crucial en la eficiente traducción de proteínas. La disminución de m⁵C38 como resultado de la inhibición de DNMT2 por señales de estrés u otros estímulos, promueve la fragmentación del ARNt, generando así 5'-tsRNAs. Esta cascada de eventos resulta en defectos en la síntesis de grupos específicos de proteínas (por ejemplo, proteínas enriquecidas en ácido aspártico), lo que a su vez conduce a irregularidades en procesos celulares fundamentales como la diferenciación. Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

A pesar de la falta de efecto sobre la traducción global, la pérdida individual de DNMT2 altera la producción de un grupo específico de proteínas (Shanmugam et al., 2015; Tuorto et al., 2015). Por un lado, la modificación m⁵C38 promueve la aminoacilación del ARNt-Asp, tanto *in vitro* como *in vivo*, facilitando la traducción de proteínas enriquecidas en ácido aspártico, cuya síntesis se encuentra significativamente reducida en ausencia de DNMT2 (Shanmugam et al., 2015). Por otro lado, la actividad catalítica de DNMT2 resulta muy relevante en el mantenimiento de la fidelidad de la traducción, ya que la metilación en C38 permite la discriminación entre codones casi homólogos, es decir, codones muy similares en secuencia pero que codifican distintos aminoácidos (por ejemplo, GAC, que codifica ácido aspártico y GAG, que codifica ácido glutámico) (Tuorto et al., 2015). De hecho, se han detectado elevados niveles de sustituciones aminoacídicas en codones reconocidos por ARNt metilados por DNMT2 en ratones mutantes para esta proteína (ratones *Dnmt2* KO) (Tuorto et al., 2015). Cabe destacar que, el aumento de la escisión del ARNt vinculado a la pérdida de DNMT2 reduce su estabilidad y dificulta la discriminación fidedigna de los codones codificados (Schaefer et al., 2010; Tuorto et al., 2015). Además, tal y como se ha mencionado

previamente, el incremento en la generación de tsRNAs en ausencia de la modificación m^5C38 o de DNMT2, también puede conducir a la inhibición del proceso de traducción (Ivanov et al., 2011; Ivanov et al., 2014; Sobala y Hutvagner, 2013).

Como consecuencia de su influencia sobre la traducción de proteínas, el papel de DNMT2 adquiere una especial relevancia durante el proceso de diferenciación celular (**Figura I13**). Por ejemplo, en ausencia de DNMT2, los ratones experimentan un retraso significativo en el proceso de osificación endocondral, clave en la formación de tejido óseo (Tuorto et al., 2015). Este fenotipo fue posteriormente asociado a una menor capacidad de diferenciación de las células troncales hematopoyéticas, causada por la reducción de la síntesis de proteínas específicas como NESTINA y PERIOSTINA (Tuorto et al., 2015). La función de DNMT2 también es necesaria para el mantenimiento de una capacidad de diferenciación equilibrada. Este es el caso de las células estromales mesenquimales, en las que la eliminación de DNMT2 provoca la estimulación e inhibición de su diferenciación hacia tejido osteogénico y adipogénico (Tuorto et al., 2015), respectivamente.

3.4.2.3 Herencia transgeneracional

La expresión y la actividad catalítica de DNMT2 son esenciales en la herencia "epigenética" mediada por ARN, consistente en la transmisión transgeneracional de fenotipos de manera independiente de la información genética codificada (Kiani et al., 2013; Zhang et al., 2018). Ejemplos de este fenómeno incluyen la transmisión de la modulación epigenética del gen *Kit*, que resulta en la alteración del color de la cola en ratones, y del gen *Sox9*, que induce un fenotipo de sobrecrecimiento (Kiani et al., 2013; Yu et al., 2021). Se observó que los ratones deficientes en *Dnmt2* no podían transmitir a la siguiente generación las modificaciones epigenéticas de los genes *Kit* y *Sox9*, así como sus fenotipos asociados, en comparación con aquellos que sí expresaban *Dnmt2* (Kiani et al., 2013; Yu et al., 2021). En ambos casos, tanto la presencia de DNMT2 como su actividad catalítica sobre ARNnc (por ejemplo, ARNm y ARN pequeño no codificante), son cruciales para la transmisión de estos fenotipos entre generaciones (Kiani et al., 2013).

De manera similar, otros estudios revelaron que la pérdida de m^5C38 causada por la ausencia de DNMT2, altera los perfiles de expresión de los tsRNAs de 30-40 nucleótidos presentes en el ARN espermático (Zhang et al., 2018), impidiendo así la transmisión de trastornos metabólicos inducidos por una dieta alta en grasas a la progenie (Chen et al., 2016; Zhang et al., 2018) (**Figura I14**).

Los tsRNAs tienen la capacidad de interactuar con otros ARNs, incluyendo ARNm y ARNnc, tanto en el compartimento espermático como después de la fertilización. Este proceso permite la regulación de cascadas transcripcionales y otras rutas que influyen en la herencia transgeneracional de caracteres (Chen et al., 2016; Schorn et al., 2017; Sharma et al., 2016). No obstante, el mecanismo molecular mediante el cual ejercen su función aún no ha sido completamente dilucidado. Por lo tanto, una exploración más exhaustiva de su función, así como de las enzimas que regulan su perfil de expresión, como es el caso de DNMT2, contribuirá a un mejor entendimiento de las circunstancias en las que un fenotipo adquirido puede ser transmitido a la descendencia.

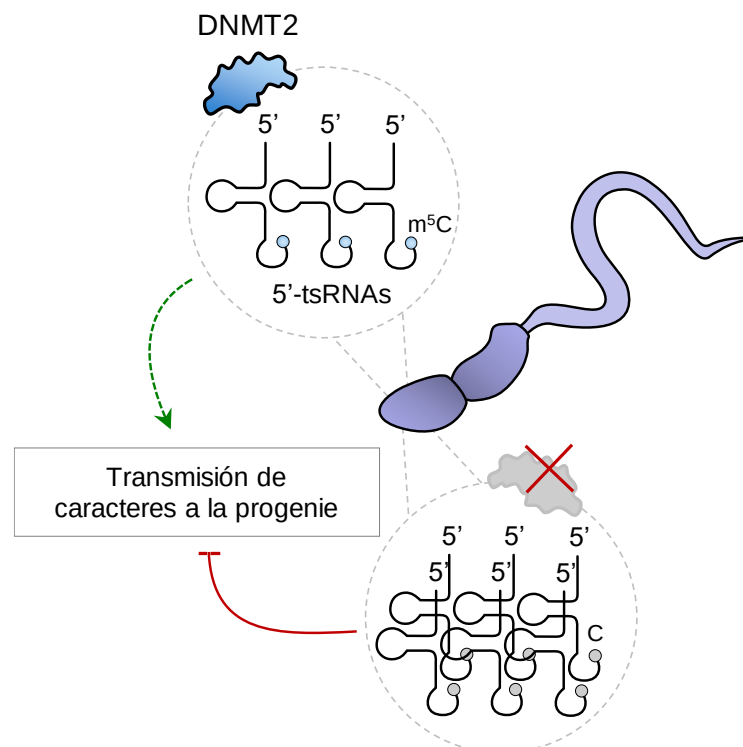


Figura I14. La actividad catalítica DNMT2 en ARN interviene en la herencia transgeneracional. Ilustración del papel central de DNMT2 en el establecimiento de perfiles de expresión y modificación de tsRNAs en el esperma, lo cual es crucial para la transmisión intergeneracional de caracteres adquiridos, como los trastornos metabólicos inducidos por una dieta alta en grasas. Esta imagen, de elaboración propia, fue creada con el programa Inkscape (v1.3).

3.4.3 Fenotipo y procesos patológicos asociados

La aparente falta de anomalías fenotípicas consistentes entre distintos modelos de estudio en ausencia de DNMT2 ha constituido un importante obstáculo en la definición de su función biológica. Los estudios iniciales indicaron que en condiciones estándar de laboratorio, los organismos sin DNMT2 (referidos en adelante como organismos *Dnmt2* KO), incluyendo ratones, moscas y plantas, eran viables, fértiles y morfológicamente indistinguibles de los organismos de tipo salvaje (WT, por sus siglas en inglés) (Goll et al., 2006). No obstante, investigaciones más detalladas revelaron defectos en el desarrollo y la diferenciación de tejidos específicos en algunos organismos modelo. Por ejemplo, la inhibición de la expresión de *Dnmt2* en embriones de pez cebra (*Danio rerio*) resultó en defectos en la diferenciación tardía y el desarrollo de órganos específicos, como el cerebro, la retina y el hígado (Rai et al., 2007). El fenotipo original pudo restaurarse mediante la expresión forzada de la forma citoplasmática y catalíticamente activa de DNMT2, pero no con una mutante catalíticamente inactiva ni con una forma nuclear (Rai et al., 2007). Esto sugiere que la presencia de la modificación m⁵C38 en los ARNt citoplasmáticos es esencial durante los procesos de desarrollo embrionario y organogénesis.

En consonancia con estas evidencias, otros estudios señalaron deficiencias en el desarrollo de ratones tras la eliminación de DNMT2, ya sea de forma individual o en combinación con NSUN2 (Tuorto et al., 2015; Tuorto et al., 2012). Es importante destacar que los ratones doble KO para *Dnmt2* y *Nsun2* mostraron un fenotipo más grave, con defectos severos en su desarrollo y diferenciación celular. Estos ratones presentaban un tamaño

considerablemente menor en comparación con el control, junto con una viabilidad significativamente reducida (Tuorto et al., 2012). Además, se observaron fallos en la diferenciación en órganos como el cerebro o el tejido adiposo. Este fenotipo se asoció con una dramática disminución de los niveles de m⁵C en el ARNt, y una concomitante reducción de la tasa de síntesis global de proteínas (Tuorto et al., 2012). Por otro lado, la eliminación individual de DNMT2 reveló un retraso en la osificación endocondral y defectos en el proceso de diferenciación durante la hematopoyesis en ratones de 8 días de edad (Tuorto et al., 2015). Este fenotipo se relacionó con la disminución de la fidelidad de la traducción durante la decodificación de los codones de ácido aspártico debido a la ausencia de m⁵C38 en el ARNt portador de este aminoácido (Tuorto et al., 2015).

Como se mencionó anteriormente, diversos estudios han señalado fenotipos dependientes del contexto, asociando la función de DNMT2 con la señalización del estrés en *Drosophila melanogaster* (Schaefer et al., 2010), con el estado nutricional en *Schizosaccharomyces pombe* (Becker et al., 2012) y la herencia transgeneracional mediada por ARN en *Mus musculus* (Kiani et al., 2013; Yu et al., 2021; Zhang et al., 2018). A pesar de que la ausencia de DNMT2 generalmente resulta en un fenotipo relativamente leve en condiciones estándar de laboratorio, las moscas deficientes en *Dnmt2* mostraron una vida significativamente acortada bajo condiciones de estrés, mientras que su sobreexpresión condujo a una mayor resistencia a los mismos estímulos (Lin et al., 2005; Schaefer et al., 2010). La sensibilidad de estos mutantes a condiciones de estrés se atribuye al aumento de la escisión de ARNt, junto con defectos en el proceso de traducción, especialmente en proteínas enriquecidas en ácido aspártico, causados por la pérdida de m⁵C38 (Shanmugam et al., 2015). En consonancia, la inhibición de *DNMT2* se ha asociado con una mayor sensibilización de las células humanas (Lewinska et al., 2018) y murinas (Lewinska et al., 2017) frente al estrés oxidativo. En conjunto, estas evidencias establecen diversas funciones biológicamente relevantes para DNMT2, sugiriendo que la modificación m⁵C38 en ARNt promueve el desarrollo y la diferenciación en distintos organismos y contextos celulares al regular la síntesis de proteínas.

La actividad de DNMT2 también desempeña un papel crucial en el mantenimiento de la integridad del genoma, aunque en este contexto la proteína actúa sobre otros tipos de ARN. Estudios han vinculado la pérdida de función de DNMT2 con una menor actividad de la enzima telomerasa y una disminución en la longitud de los telómeros, así como con un mayor daño en el ADN (Chen et al., 2020; Lewinska et al., 2017). Específicamente, el silenciamiento de DNMT2 conduce a una reparación ineficaz de las roturas de doble cadena (DSBs, por sus siglas en inglés) debido a defectos en el proceso de recombinación homóloga (HR, por sus siglas en inglés) (Chen et al., 2020; Yang et al., 2022; Zhu et al., 2021). En estas regiones, DNMT2 deposita m⁵C sobre ARNm en híbridos de ADN y ARN promoviendo así la HR para su reparación (Chen et al., 2020; Keskin et al., 2014; Yang et al., 2022; Zhu et al., 2021). Por lo tanto, la actividad catalítica de DNMT2 en ARN desempeña un papel importante en la supervivencia celular después del daño en el ADN.

Además, la eliminación de DNMT2 también está asociada con la desrepresión de elementos transponibles (TEs, por sus siglas en inglés) en el genoma de varios organismos (Betlej et al., 2022; Genencher et al., 2018; Phalke et al., 2009). En el caso de *Drosophila melanogaster*, la expresión de estos elementos pudo ser restaurada mediante la expresión forzada de una forma catalíticamente inactiva de la proteína, sugiriendo que DNMT2 podría regular los niveles de estos elementos de manera independiente de la metilación del ARN

(Genenncher et al., 2018). Sin embargo, hasta el momento, se desconoce si DNMT2 presenta actividad catalítica directa sobre los transcritos de TEs.

Otro fenotipo asociado con la ausencia de DNMT2 es la hipertrofia cardíaca, detectada en ratones adultos (Ghanbarian et al., 2016). En este caso, la eliminación de esta proteína resulta en una reducción significativa de la metilación del ARN no codificante *Rn7sk*, que en condiciones normales se asocia con el factor regulador de la elongación de la transcripción, P-TEFb, para inhibir la activación de la ARN polimerasa II. De este modo, un complejo P-TEFb hiperactivo conduce a un aumento en la transcripción y como resultado, a un sobrecrecimiento del corazón en estos organismos (Ghanbarian et al., 2016). Estos hallazgos también fueron recapitulados en ESCs sugiriendo que la deposición de m⁵C mediada por DNMT2 podría regular la interacción entre *Rnsk7* y P-TEFb (Ghanbarian et al., 2016). De manera interesante, DNMT2 también fue identificada en células de leucemia humana, junto con las proteínas de unión a ARN NSUN3 y hnRNPK, formando un complejo con P-TEFb (Cheng et al., 2018). Este complejo recluta a la ARN polimerasa II en el ARN naciente, dando lugar a la formación de una estructura de cromatina sensible a la 5-azacitidina (5-AZA), un agente hipometilante del ADN y ARN comúnmente utilizado en el tratamiento de neoplasias hematológicas (Cheng et al., 2018). Sin embargo, todavía se desconoce el papel específico de DNMT2 en este contexto.

Recientemente, la actividad de DNMT2 también ha sido relacionada con la formación de tumores. Su expresión se encuentra consistentemente desregulada en cientos de muestras tumorales y se han identificado más de 90 mutaciones somáticas en este gen en varios tipos de cáncer (Elhardt et al., 2015; Forbes et al., 2015). Aunque la investigación sobre la función de DNMT2 en cáncer es actualmente limitada, existen evidencias emergentes de su participación en la regulación biológica de ciertos tipos celulares tumorales. En tumores gástricos, por ejemplo, se observa una tendencia al alza en la regulación de *DNMT2*, de modo que su expresión elevada correlaciona con un peor pronóstico para los pacientes (Sun et al., 2021). Además, DNMT2 participa en la regulación de procesos clave en el crecimiento tumoral, como la proliferación, migración e invasión celulares. En células de cáncer gastrointestinal, la inhibición de DNMT2 resulta en una reducción de la capacidad de migración e invasión, mientras que en células HEK293T, estas capacidades se ven reducidas tras la inhibición de la metilación del ARNm *UGP2*, dependiente de DNMT2 (Xue et al., 2019). En fibroblastos humanos, el silenciamiento de DNMT2 promueve la regulación al alza de miRNAs reguladores de la proliferación, algunos de ellos implicados en la supresión de tumores, como miR-28-3p, miR-34a-3p y miR-30b-5p, entre otros (Lewinska et al., 2018). Además, en algunos tipos celulares, como células de osteosarcoma, se ha detectado un aumento en la actividad de TEs después de la eliminación de DNMT2, lo que constituye otra evidencia de su papel importante en el contexto del cáncer (Betlej et al., 2022).

Por otro lado, una menor expresión de *DNMT2* correlaciona con una menor resistencia a fármacos antitumorales en algunos tipos de cáncer (Bloniarz et al., 2021; Chen et al., 2020; Zhu et al., 2021). La pérdida de la actividad catalítica de DNMT2 en células de cáncer de mama, ovario y osteosarcoma contribuye a aumentar la sensibilidad al daño celular, ya que algunos mecanismos de reparación del ADN, como la HR, se encuentran comprometidos (Betlej et al., 2022; Chen et al., 2020; Zhu et al., 2021). En tumores de mama y en células de osteosarcoma deficientes en DNMT2, se ha observado una mayor sensibilidad a los inhibidores de PARP1, como olaparib (Chen et al., 2020). Además, las células de osteosarcoma también mostraron una mayor sensibilidad frente a otros tratamientos antitumorales como el etopósido o el cisplatino (Betlej et al., 2022; Zhu et al., 2021). También, tal y como se ha mencionado con anterioridad,

DNMT2 participa en la resistencia a 5-AZA en células de leucemia (Cheng et al., 2018). Análogamente, en xenoinjertos murinos derivados de células humanas de cáncer de ovario deficientes en DNMT2, el tamaño de los tumores se reduce significativamente tras el tratamiento con cisplatino (Zhu et al., 2021). Además, la menor expresión de DNMT2 se asocia a una mejor respuesta a la radioterapia en pacientes con cáncer de mama (Chen et al., 2020). Sin embargo, en glioblastoma, la eliminación del gen *DNMT2* puede promover la resistencia de estas células a terapias antitumorales, observándose una disminución de las células apoptóticas y senescentes, así como de la autofagia (Bloniarz et al., 2021). Estas discrepancias subrayan la complejidad de la regulación de la biología tumoral por parte de alteraciones en DNMT2, enfatizando la necesidad de investigaciones adicionales para comprender su papel dual en la progresión del cáncer y, por ende, identificar potenciales estrategias terapéuticas basadas en su regulación.

En resumen, aunque inicialmente las deficiencias en la función de DNMT2 fueron consideradas poco relevantes, estudios más recientes han destacado su crucial implicación en el correcto transcurso de procesos como el desarrollo embrionario, la diferenciación y organogénesis, la reparación de daños y respuesta al estrés, así como la herencia transgeneracional de caracteres (**Tabla 2**). Además, debido a su importancia en la homeostasis de tejidos, los defectos funcionales de DNMT2 han sido relacionados con el inicio y progresión de distintos tipos de cáncer. Es relevante señalar que, dado que la expresión de DNMT2 varía notablemente entre líneas celulares de cánceres humanos (Schaefer et al., 2009), no es sorprendente que en algunos casos actúe como supresor de tumores, mientras que en otros promueva la formación de estos. Sin embargo, aún se desconoce cómo estas variaciones en su función influyen en el proceso de carcinogénesis y si dependen o no de su actividad catalítica sobre el ARN.

Genotipo	Organismo	Fenotipo asociado	Referencia
<i>Dnmt2</i> KO	<i>Mus musculus</i>	Pérdida de la herencia transgeneracional de caracteres adquiridos	(Kiani et al., 2013; Yu et al., 2021; Zhang et al., 2018)
		Retraso en la osificación endocondral	(Tuorto et al., 2015)
		Defectos en la diferenciación hematopoyética y en el tejido adiposo	(Tuorto et al., 2015)
		Hipertrofia cardíaca	(Ghanbarian et al., 2016)
	<i>Drosophila melanogaster</i>	Defectos en la síntesis de proteínas ricas en Asp	(Shanmugam et al., 2015; Tuorto et al., 2015)
		Menor vida media bajo condiciones de estrés	(Schaefer et al., 2010)
<i>Dnmt2</i> KD	<i>Homo sapiens</i>	Desrepresión de TEs	(Genenncher et al., 2018; Phalke et al., 2009)
		Mayor sensibilidad celular bajo estrés oxidativo	(Betlej et al., 2022; Lewinska et al., 2018)
		Cambios en la respuesta a fármacos antitumorales	(Bloniarz et al., 2021)
	<i>Danio rerio</i>	Defectos en la reparación de DSBs por HR	(Betlej et al., 2022; Chen et al., 2020)
		Desrepresión de TEs	(Betlej et al., 2022)
		Defectos en diferenciación (cerebro, retina, hígado)	(Rai et al., 2007)
<i>Dnmt2</i> KD	<i>Mus musculus</i>	Menor longitud de telómeros	(Lewinska et al., 2017)
	<i>Homo sapiens</i>	Reducida migración y proliferación en células de cáncer gastrointestinal	(Sun et al., 2021)
		Cambios en la respuesta a fármacos antitumorales	(Chen et al., 2020; Zhu et al., 2021)
<i>Dnmt2/ Nsun2</i> DKO	<i>Mus musculus</i>	Elevada tasa de muerte perinatal Fenotipo subdesarrollado Menor síntesis global de proteínas	(Tuorto et al., 2012)

Tabla 2. Fenotipos asociados a la pérdida de función de DNMT2. Resumen de los principales fenotipos vinculados a la eliminación (KO) o inhibición (KD) individual de DNMT2, así como a la eliminación conjunta con NSUN2 (DKO); se especifican los organismos en los cuales han sido descritos.

4. OBJETIVOS

Aunque inicialmente se consideraban las deficiencias en la función de DNMT2 como poco relevantes, investigaciones más recientes han resaltado su implicación crucial en procesos como el desarrollo embrionario y la diferenciación celular, entre otros. Además, los defectos funcionales de DNMT2 se han relacionado con el inicio y la progresión de diversos tipos de cáncer, debido a su papel fundamental en la homeostasis de tejidos.

En este contexto, el presente proyecto de tesis doctoral se enfocará en someter a prueba la hipótesis de que *DNMT2 desempeña funciones reguladoras importantes tanto en el control de la pluripotencia como en el desarrollo del cáncer*. Por ende, el objetivo central será abordar la relevancia funcional de DNMT2 en los procesos de adquisición, mantenimiento y salida de pluripotencia, así como en la biología de las células cancerosas. Para ello, se proponen los siguientes objetivos específicos:

- **Objetivo 1:** Dilucidar la función de la proteína DNMT2 en el proceso de reprogramación celular somática.
- **Objetivo 2:** Determinar el papel de DNMT2 en el mantenimiento y salida del estado celular pluripotente.
- **Objetivo 3:** Explorar nuevas funciones de la proteína DNMT2 en contextos celulares asociados al cáncer.

Se espera que los hallazgos obtenidos en torno a estos objetivos contribuyan a un mejor conocimiento de las bases de la pluripotencia y los fundamentos moleculares de la reprogramación celular somática. Además, los resultados obtenidos tienen el potencial de generar información valiosa para superar las limitaciones actuales en el uso terapéutico de células pluripotentes, así como para avanzar en la comprensión de los mecanismos que subyacen a la participación de DNMT2 en procesos cancerosos.

5. MATERIALES Y MÉTODOS

5.1 REACTIVOS Y KITS

Todos los reactivos usados en la tesis fueron de calidad de investigación y se detallan en la **Tabla suplementaria 1**, con sus concentraciones específicas en las secciones correspondientes. Los kits y el material utilizado están enumerados en las **Tablas suplementarias 2 y 3**, respectivamente.

5.2 MODELOS EXPERIMENTALES

Los tipos celulares utilizados en esta tesis se describen a continuación:

Nombre	Fuente	Identificación
CJ7 (ESCs de ratón)	Dr. Jianlong Wang (Universidad de Columbia, EE. UU.)	RRID:CVCL_C316
<i>Dnmt2</i> KI (ESCs de ratón)	Este estudio	n/a
<i>Dnmt2</i> KO (ESCs de ratón)	Este estudio	n/a
<i>Dnmt2</i> KO ^{+WT} (ESCs de ratón)	Este estudio	n/a
<i>Dnmt2</i> KO ^{+C79A} (ESCs de ratón)	Este estudio	n/a
<i>Dnmt2</i> KO ^{+G155V} (ESCs de ratón)	Este estudio	n/a
<i>Dnmt2</i> KO ^{+R371H} (ESCs de ratón)	Este estudio	n/a
TNGA (pre-iPSCs de ratón)	Dr. Jose C.R. Silva (Universidad de Cambridge, R. U.)	n/a
O4G MEFs	Dr. Jianlong Wang (Universidad de Columbia, EE. UU.)	n/a
CD34/H2BGFP MEFs	Dr. Carlos-Filipe Pereira (Universidad de Lund, Suecia)	n/a
Fibroblastos adultos de ratón	Este estudio	n/a
HDFs (fibroblastos humanos)	Dr. Carlos-Filipe Pereira (Universidad de Lund, Suecia)	n/a
MCF-7	Dr. Román Pérez (Universidad de Santiago de Compostela, España)	RRID:CVCL_0031
MDA-MB-231	Dr. Román Pérez (Universidad de Santiago de Compostela, España)	RRID:CVCL_0062
HEK293T	ATCC	RRID:CVCL_0063
Plat-E	Cell Biolabs	RRID:CVCL_B488

- **CJ7 ESCs:** células troncales embrionarias derivadas de la masa celular interna del blastocisto preimplantacional de embriones macho de ratón 129S1/SvImJ.
- ***Dnmt2* KI ESCs:** células CJ7 *knockin* (KI) en las cuales se ha insertado una etiqueta endógena 3xFLAG en el extremo C-terminal de la región codificante del gen *Dnmt2*, utilizando la tecnología CRISPR-Cas9 (consultar las secciones 5.4.8 y 5.5).
- ***Dnmt2* KO:** células CJ7 *knockout* (KO) para el gen *Dnmt2*, generadas mediante la tecnología CRISPR-Cas9 (consultar las secciones 5.4.7 y 5.5).
- ***Dnmt2* KO^{+WT}, +C79A, +G155V o +R371H:** células pluripotentes *Dnmt2* KO en las que se ha inducido la expresión forzada de la proteína DNMT2 humana (WT) o de versiones mutantes catalíticas (C79A, G155V o R371H) mediante la transfección con vectores PiggyBac (PB) (ver secciones 5.4.4 y 5.5).
- **TNGA pre-iPSCs:** células derivadas de intermediarios de reprogramación avanzados (pre-iPSCs) que portan un constructo IRES-EGFP en el extremo C-terminal del gen *Nanog*. Estas células se generaron mediante la reprogramación de fibroblastos de ratón, induciendo la expresión forzada de los factores de reprogramación OCT4, SOX2, KLF4 y c-MYC

(OSKM) a través de la transducción con vectores retrovirales pMX, seguida de la selección de células negativas para *Nanog*-GFP mediante citometría de flujo (ver sección 5.15.2).

- **MEFs O4G:** fibroblastos derivados de embriones de ratón en el día E13.5 tras fecundación, con un fondo genético resultante del cruce entre las cepas C57BL/6 y 129S4/SvJae. Estos fibroblastos portan un constructo IRES-EGFP situado en el extremo C-terminal del gen *Oct4*.
- **CD34/H2BGFP MEFs:** Fibroblastos embrionarios obtenidos de embriones de ratón en el día E14.5 tras fecundación, con un fondo genético C57BL/6, que portan un constructo H2B-GFP bajo el control del promotor del gen *CD34* humano.
- **Fibroblastos adultos de ratón:** fibroblastos primarios obtenidos de las orejas de ratones macho, con un fondo genético C57BL/6J, de 5 meses de edad, criados bajo condiciones de dieta estándar.
- **HDFs:** fibroblastos humanos obtenidos del tejido dérmico de individuos adultos (Science Cell, 2320).
- **MCF-7:** línea celular derivada de un adenocarcinoma de mama humano perteneciente a una paciente de etnia blanca de 69 años.
- **MDA-MB-231:** línea celular derivada de un adenocarcinoma de mama humano perteneciente a una paciente de etnia blanca de 40 años.
- **HEK293T:** células inmortalizadas de riñón embrionario humano, transfectadas con el antígeno T SV40 para posibilitar la propagación de vectores que contienen la región replicativa SV40. A lo largo de este estudio, se utilizó esta línea celular para generar partículas lentivirales que portaran shRNAs clonados en el vector pLKO, así como constructos destinados a la sobreexpresión de genes clonados en los vectores FUW y FUW-Tet-On (ver secciones 5.4.5 y 5.7.1).
- **Plat-E:** línea celular derivada de HEK293T que expresa los genes *gag*, *pol* y *env*, facilitando el empaquetamiento retroviral a través de la transfección de un solo plásmido. Esta línea celular se utilizó para producir partículas retrovirales portando constructos destinados a la sobreexpresión de genes clonados en el vector pMX (ver apartados 5.4.6 y 5.7.1).

5.3 CULTIVO CELULAR

Cada tipo celular utilizado en este estudio, junto con el material destinado a su cultivo, se mantuvo en condiciones de esterilidad y fue manipulado en cabinas de flujo laminar.

A lo largo de la investigación, las células fueron cultivadas en presencia de suero fetal bovino (FBS; Gibco, 10270106) previamente inactivado a 56°C durante 30 minutos. En caso necesario, se llevó a cabo la individualización celular mediante tripsina-EDTA (Sigma-Aldrich, T3924). Los cultivos celulares se mantuvieron a 37°C con un 5% de CO₂.

Las ESCs CJ7 y las pre-iPSCs TNGA fueron cultivadas en placas de cultivo de poliestireno previamente tratadas con gelatina (Sigma-Aldrich, G9391) al 0,1%, diluida en agua ultrapura (Invitrogen, 10977-035) y estéril. Las placas se incubaron con gelatina durante al menos 30 minutos a 37°C y 5% de CO₂. La gelatina fue retirada justo antes de la siembra de células. Estos tipos celulares se mantuvieron en un medio de cultivo estándar para células

pluripotentes, denominado en adelante medio pluripotente con suero y LIF (SL). Este consta de medio Eagle modificado de Dulbecco alto en glucosa (DMEM; Corning, 10-013-CV), al cual se le añadió un 15% de FBS. Además, se suplementó con un 1% de mezcla de nucleósidos en agua ultrapura [adenosina (Sigma-Aldrich, A4036), citidina (Sigma-Aldrich, C4654), guanosina (Sigma-Aldrich, G6264), timidina (Sigma-Aldrich, T1895) y uridina (Sigma-Aldrich, U3003)], 200 mM de L-Glutamina (Gibco, 25030-081), 100 mM de aminoácidos no esenciales (Gibco, 11140-050), 100 U/mL de penicilina/estreptomicina (Gibco, 15070-063), 0,1 mM de β -mercaptoetanol (Sigma-Aldrich, M-6250) y 1.000 U/mL de factor inhibidor de leucemia (LIF; producido y probado en el laboratorio).

Los fibroblastos, HEK293T, Plat-E y las líneas tumorales se cultivaron directamente en placas de cultivo celular de poliestireno y se mantuvieron en medio completo, compuesto por DMEM alto en glucosa, suplementado con un 10% de FBS, 200 mM de L-Glutamina y 100 U/mL de penicilina/estreptomicina.

Con el fin de preservar las células a largo plazo, se procedió a su individualización y recolección mediante centrifugación a 300 xg durante 5 minutos a temperatura ambiente (RT). Tras retirar el sobrenadante, las células se resuspendieron en un medio de congelación compuesto por un 90% del medio de cultivo correspondiente en cada caso y un 10% de dimetilsulfóxido (DMSO; Sigma-Aldrich, 75-21-8). Las células se almacenaron a -80°C durante un máximo de 48 horas y posteriormente fueron trasladadas a tanques de nitrógeno líquido.

5.4 GENERACIÓN DE VECTORES PLASMÍDICOS

5.4.1 Producción de bacterias competentes

En primer lugar, se sembraron células de *Escherichia coli* de las cepas DH5 α o Stbl3 (esta última se usó para vectores virales o sistemas de transposasa con el objetivo de minimizar las posibilidades de recombinación) en placas de agar Luria-Bertani (LB) sin antibiótico mediante el método de siembra por agotamiento. Posteriormente, se incubaron durante 16 horas a 37°C. Las colonias individuales fueron transferidas a 25 mL de medio líquido LB y se incubaron durante otras 16 horas a 37°C con una velocidad de 225 revoluciones por minuto (rpm).

Luego, se inoculó 1 mL de la solución de crecimiento bacteriano en 100 mL de medio LB, el cual se incubó a 37°C y 225 rpm hasta que la densidad óptica del cultivo alcanzó un valor de 0,55 de absorbancia a 600 nm. A continuación, el cultivo bacteriano se enfrió en hielo durante 10 minutos, y posteriormente se centrifugó a 2.500 xg durante 5 minutos a 4°C. El sedimento bacteriano resultante se resuspendió en tampón de transformación Inoue, previamente enfriado, compuesto por 55 mM de MnCl₂·4H₂O (Sigma-Aldrich, M8266), 15 mM de CaCl₂·2H₂O (Fisher Chemical, VL2151000), 250 mM de KCl (Fisher Chemical, P/4240/60) y 10 mM de PIPES pH 6.7 (Sigma-Aldrich, P6757) en agua ultrapura (Invitrogen, 10977-035). Luego, las bacterias fueron centrifugadas nuevamente a 2.500 xg durante 5 minutos a 4°C.

A continuación, las bacterias sedimentadas se resuspendieron en tampón de transformación Inoue suplementado con DMSO (Sigma-Aldrich, D8418) a una concentración final del 7% y se incubaron en hielo durante 10 minutos. Finalmente, la solución de bacterias competentes se distribuyó en alícuotas de 100 μ L, las cuales fueron rápidamente congeladas en un baño de hielo seco y etanol. Para su conservación a largo plazo, se mantuvieron a -80°C.

5.4.2 Transformación de bacterias competentes

Con el fin de amplificar los vectores plasmídicos generados, se llevó a cabo la transformación de bacterias competentes. Para ello, las alícuotas de bacterias competentes fueron descongeladas en hielo, seguido de la adición de un máximo de 10 μ L de ADN plasmídico por cada 100 μ L de bacterias. Esta mezcla se incubó durante 25 minutos en hielo, después de lo cual se aplicó un golpe de calor durante 45 segundos a 42°C, seguido de una incubación en hielo de 5 minutos. Posteriormente, se añadieron 300 μ L de medio SOC (compuesto por 20 g de triptona, 5 g de extracto de levadura y 0,5 g de NaCl en 1 L de agua), y la mezcla se incubó durante 1 hora a 37°C y 225 rpm. Finalmente, se sembraron 100-200 μ L del cultivo bacteriano en placas de agar LB con antibiótico, que se incubaron durante 16 horas a 37°C.

5.4.3 Clonación de shRNAs en pLKO

Con el fin de silenciar la expresión del gen humano o murino que codifica para DNMT2, se utilizaron ARNs de horquilla corta (shRNAs), los cuales se clonaron en el vector de expresión lentiviral pLKO.1 (Addgene, 8453) con resistencia a ampicilina y puromicina, lo que permite su selección en células procariotas y eucariotas, respectivamente. En primer lugar, se diseñaron oligonucleótidos dirigidos contra secuencias específicas de los genes humanos y murinos que codifican para DNMT2, y del gen *Luciferasa* como control, mediante el portal web del Broad Institute GPP (<https://portals.broadinstitute.org/gpp/public/>). Cada par de oligonucleótidos fue hibridado en una reacción compuesta por 2,5 μ M de oligonucleótido directo, 2,5 μ M de oligonucleótido inverso y 1X de tampón de reacción NEB (New England BioLabs, B7002S). La reacción se incubó durante 10 minutos a 95°C, luego 10 minutos a 70°C, y finalmente se enfrió a una velocidad de 1°C por cada 2,5 minutos hasta llegar a 25°C, estabilizándose finalmente a 4°C en un termociclador.

Estos oligonucleótidos se diseñaron para generar extremos cohesivos complementarios a las digestiones producidas por las enzimas de restricción *AgeI* (Thermo Scientific, ER1461) y *EcoRI* (Thermo Scientific, ER0271). Posteriormente, el vector fue digerido con las enzimas *AgeI* y *EcoRI*, y el producto se separó en un gel de agarosa mediante electroforesis y se purificó con el kit GenElute Gel Extraction (Sigma-Aldrich, NA1111), siguiendo las instrucciones del fabricante. Luego, se llevó a cabo la ligación del vector linealizado con los oligonucleótidos hibridados mediante la ligasa de ADN T4 (Thermo Scientific, EL0016), siguiendo las instrucciones del fabricante. Cada constructo generado se verificó mediante secuenciación de Sanger.

Las secuencias de los oligonucleótidos empleados se pueden consultar en las **Tablas suplementarias 5 y 11**.

5.4.4 Clonación de regiones codificantes en PiggyBac

Para investigar la influencia de la actividad catalítica de DNMT2 en la regulación del estado celular pluripotente, se utilizó el sistema de transposones PiggyBac (PB) para expresar variantes WT o mutantes catalíticas de DNMT2 con tres etiquetas Flag (3xFLAG) en el extremo N-terminal (Katter et al., 2013). Inicialmente, se partió de plásmidos pet28a(+) que contenían las secuencias codificantes de las proteínas humanas DNMT2 WT y sus mutantes (C79A, G155V, R371H), proporcionados por el laboratorio del Dr. Albert Jeltsch de la Universidad de Stuttgart. Las mutaciones C79A, G155V y R371H se han documentado previamente en diversas publicaciones del grupo (Elhardt et al., 2015; Jurkowski et al., 2008) como causantes de la

ausencia (C79A) o reducción (G155V, R371H) de la actividad catalítica de las proteínas mutantes DNMT2.

Posteriormente, mediante el uso de una mezcla maestra de polimerasa de alta fidelidad Phusion Green Hot Start II (Thermo Scientific, F566S) se procedió a la amplificación por PCR de las secuencias codificantes para las proteínas humanas WT y mutantes, empleando cebadores que contienen los sitios de reconocimiento de las enzimas de restricción *NheI* (Thermo Scientific, ER0971) y *PacI* (Thermo Scientific, ER2202) y siguiendo las instrucciones del fabricante. Las secuencias de los cebadores empleados se proporcionan en la **Tabla suplementaria 6**. Los productos de amplificación fueron visualizados tras una electroforesis en un gel de agarosa, del cual fueron posteriormente purificados empleando el kit GenElute Gel Extraction (Sigma-Aldrich, NA1111), siguiendo las instrucciones del fabricante.

Luego, las secuencias purificadas se ligaron de forma independiente en vectores pJET1.2 utilizando el kit CloneJet PCR (Thermo Scientific, K1231), siguiendo las instrucciones del fabricante. Los vectores generados se amplificaron mediante transformación de bacterias competentes y se purificaron con el kit GenElute Plasmid Miniprep (Sigma-Aldrich, PLN350), siguiendo las instrucciones del fabricante. Las identidades de los vectores se confirmaron mediante secuenciación de Sanger.

A continuación, los vectores pJET1.2 conteniendo las secuencias codificantes de DNMT2 WT o mutantes, así como el vector pPB-CAG-3xFLAG-pA-pgk-hph (Addgene, 48753), se digirieron con las enzimas de restricción *NheI* y *PacI*. Los fragmentos linealizados con tamaños esperados se purificaron de un gel de agarosa y se extrajeron con el kit GenElute Gel Extraction, siguiendo las instrucciones del fabricante. Luego, las secuencias de interés se clonaron en el vector linealizado utilizando la ligasa de ADN T4 (Thermo Scientific, EL0016), siguiendo las instrucciones del fabricante. Este vector posee genes de resistencia a ampicilina e higromicina, lo que permite su selección en células procariotas y eucariotas, respectivamente.

Las secuencias de los oligonucleótidos empleados para la secuenciación Sanger se pueden consultar en la **Tabla suplementaria 11**.

5.4.5 Clonación de regiones codificantes en pFUW-Tet-On

Con el propósito de investigar la dependencia de la actividad catalítica de DNMT2 en los contextos de reprogramación celular y cáncer, se utilizaron sistemas de expresión inducibles rtTA/Tet-On basados en los vectores lentivirales FUW-rtTA (Addgene, 20342) y FUW-Tet-On (Addgene, 125598), dependientes de la presencia del antibiótico doxiciclina (DOX).

En primer lugar, se procedió a la amplificación por PCR de la secuencia codificante (CDS) de *Dnmt2*^{WT} (ratón), *DNMT2*^{WT} (humano) y *DNMT2*^{C79A} (humano) a partir de vectores PB que contenían estas secuencias, empleando una mezcla maestra de polimerasa de alta fidelidad Platinum SuperFi II (Invitrogen, 12368010) y siguiendo las instrucciones del fabricante. Los cebadores empleados contenían dianas de restricción para la enzima *EcoRI*, cuyas secuencias se detallan en la **Tabla suplementaria 6** (ver sección 9). Los productos de amplificación fueron visualizados tras una electroforesis en un gel de agarosa, del cual fueron purificados empleando el kit GenElute Gel Extraction (Sigma-Aldrich, NA1111), siguiendo las instrucciones del fabricante.



Después de la purificación, las secuencias de interés se ligaron de forma independiente en vectores pJET1.2 utilizando el kit CloneJet PCR (Thermo Scientific, K1231), según las

recomendaciones del fabricante. Los vectores generados se amplificaron mediante transformación de bacterias *Stbl3* competentes y se purificaron con el kit GenElute Plasmid Miniprep (Sigma-Aldrich, PLN350), siguiendo las indicaciones del fabricante. Las identidades de los vectores se confirmaron mediante secuenciación de Sanger (secuencias de oligonucleótidos disponibles en la **Tabla suplementaria 11**).

A continuación, tanto los vectores pJET1.2 con las secuencias de interés como los vectores pFUW-Tet-On, fueron digeridos con la enzima *EcoRI* (Thermo Scientific, ER0271), siguiendo las indicaciones del fabricante. Además, el vector de destino fue desfosforilado para evitar su religación mediante el uso de fosfatasa alcalina termosensible FastAP (Thermo Scientific, EF0654), según las instrucciones del fabricante. Luego, el vector linealizado y las secuencias de interés se purificaron mediante electroforesis y extracción de gel con el kit GenElute Gel Extraction (Sigma-Aldrich, NA1111), siguiendo las indicaciones del fabricante.

Posteriormente, las secuencias codificantes se clonaron en el vector pFUW-Tet-On mediante el uso de la ligasa de ADN T4 (Thermo Scientific, EL0016), siguiendo las instrucciones del fabricante. Este vector contiene genes de resistencia a ampicilina y bleomicina, lo que permite su selección en células procariotas y eucariotas, respectivamente.

5.4.6 Clonación de regiones codificantes en pMX

Con el fin de determinar la influencia de la actividad catalítica de DNMT2 en los procesos de reprogramación celular, se utilizó el sistema de expresión retroviral constitutivo pMX. Realizamos la clonación de las secuencias codificantes de las proteínas humanas DNMT2^{WT} y DNMT2^{C79A} obtenidas a partir de los vectores PB generados en la sección 5.4.4, en el vector 3xFLAG-dCas9/pMXs-neo (Addgene, 51260), tras la eliminación de la secuencia codificante para la proteína dCas9. Inicialmente, procedimos con la digestión del vector de destino y de los vectores portadores de las secuencias codificantes utilizando las enzimas de restricción *ClaI* (Thermo Scientific, 10839530) y *NotI* (Thermo Scientific, 10397799), según las recomendaciones del fabricante. A continuación, verificamos los resultados de la digestión mediante electroforesis en gel de agarosa y purificamos los fragmentos de interés mediante el kit GenElute Gel Extraction (Sigma-Aldrich, NA1111), siguiendo las instrucciones del fabricante. Posteriormente, las secuencias codificantes se clonaron en el vector 3xFLAG-dCas9/pMXs-neo linealizado mediante el uso de la ligasa de ADN T4 (Thermo Scientific, EL0016), siguiendo las instrucciones del fabricante. Este vector contiene genes de resistencia a ampicilina y neomicina, lo que permite su selección en células procariotas y eucariotas, respectivamente.

5.4.7 Clonación de guías de ARN para la generación de ESCs *knockout* (KO) para *Dnmt2*

Para generar una línea de ESCs KO para *Dnmt2*, se utilizó la tecnología de edición génica CRISPR/Cas9 (Ran et al., 2013). Para esto, se diseñaron guías de ARN específicas contra la secuencia diana ubicada en el exón 1 del gen *Dnmt2* mediante la herramienta en línea CRISPOR (<http://crispor.tefor.net>) (Concordet y Haeussler, 2018). La guía seleccionada fue la mejor valorada por el programa, con la menor probabilidad de unión inespecífica a otras regiones.

A continuación, los oligonucleótidos complementarios de la guía fueron hibridados y fosforilados en una reacción con 2,5 μ M de oligonucleótido directo, 2,5 μ M de oligonucleótido inverso, 1X de tampón de reacción A del enzima polinucleótido quinasa T4 (PNK) (Thermo

Scientific, EK0032), 1 mM de ATP pH 7 (TCI Chemicals, A0157) y 10 unidades de T4 PNK (Thermo Scientific, EK0032). La reacción se llevó a cabo durante 30 minutos a 37°C, 5 minutos a 95°C y se estabilizó a 4°C. Las secuencias complementarias de la guía fueron diseñadas para generar extremos cohesivos que se unen a las digestiones generadas por la enzima de restricción *BbsI* (Thermo scientific, ER1012).

El vector SpCas9-2A-Puro V2.0 (Addgene, 62988) fue digerido y desfosforilado con *BbsI* y fosfatasa alcalina termosensible FastAP (Thermo Scientific, EF0654), siguiendo las instrucciones del fabricante. El vector linealizado se purificó mediante electroforesis en gel de agarosa y el kit GenElute Gel Extraction (Sigma-Aldrich, NA1111), siguiendo las instrucciones del fabricante. Posteriormente, se llevó a cabo la ligación del vector linealizado y la guía de ARN utilizando la ligasa de ADN T4 (Thermo Scientific, EL0016), siguiendo las recomendaciones del fabricante. La secuencia del vector generado fue validada por secuenciación de Sanger. Es importante destacar que este vector contiene genes de resistencia a ampicilina y puromicina, lo que permite su selección en células procariotas y eucariotas, respectivamente. Las secuencias de la guía de ARN y del cebador empleado para la secuenciación están disponibles en las **Tablas suplementarias 4 y 11** del apéndice, respectivamente (ver sección 9).

5.4.8 Clonación de guías de ARN para la generación de ESCs *knockin* (KI) para *Dnmt2*

Con el objetivo de generar una línea homocigótica de ESCs KI que llevara la secuencia de codificación para el péptido FLAG en el extremo C-terminal de *Dnmt2*, se empleó la tecnología de edición génica CRISPR/Cas9, junto con el vector donador pFETCh-3xFlag-P2A-NeoR (Addgene, 63934) (Savic et al., 2015).

Para esto, se llevó a cabo la clonación de dos regiones homólogas al gen *Dnmt2*, de 700 y 748 pares de bases (pb), ubicadas en proximidad al codón de parada. Estas regiones se clonaron en el vector pFETCh-3xFlag-P2A-NeoR (Addgene, 63934), que contiene tres secuencias FLAG (3xFLAG), así como genes de resistencia a kanamicina y neomicina para la selección en células procariotas y eucariotas, respectivamente. Se añadieron colas a estas regiones para facilitar el ensamblaje mediante Gibson (Gibson et al., 2009).

La transfección del plásmido que albergaba estas regiones permitió la inserción del constructo de interés por recombinación homóloga en el ADN endógeno, tras el corte generado con el sistema CRISPR/Cas9.

Para la amplificación de las dos regiones del gen *Dnmt2*, de 748 pb, aguas arriba; y de 700 pb, aguas abajo del codón de parada, se utilizó la mezcla maestra de polimerasa Platinum SuperFi II PCR (Invitrogen, 12368010), siguiendo las instrucciones del fabricante. Estas regiones fueron amplificadas a partir de ADN de ESCs CJ7. Los cebadores utilizados se detallan en la **Tabla suplementaria 7** y fueron diseñados para generar extremos cohesivos compatibles con los cortes producidos por las enzimas de restricción *BsaI* (Thermo Scientific, ER0291) y *BbsI* (Thermo Scientific, ER1012).

Tras la digestión del vector de destino, los productos de PCR con las secuencias homólogas al gen *Dnmt2* fueron purificados mediante el kit GenElute PCR Clean-up (Sigma-Aldrich, NA1020-1KT), siguiendo las instrucciones del fabricante. Finalmente, la clonación de las secuencias de interés en el vector de destino se llevó a cabo mediante la ligación con el kit GeneArt™ Gibson Assembly HiFi (Invitrogen, A46627), siguiendo las recomendaciones del

fabricante. La identidad del plásmido resultante se validó mediante secuenciación de Sanger. Las secuencias de oligonucleótidos empleadas para la validación de la identidad del vector final se encuentran disponibles en la **Tabla suplementaria 11**.

5.4.9 Amplificación y purificación de ADN plasmídico

Para purificar el ADN plasmídico amplificado a través de la transformación de bacterias competentes, las colonias obtenidas en placas de agar LB con antibiótico se cultivaron de manera individual en 8 mL de medio líquido LB. Este medio estaba suplementado con 100 µg/mL de ampicilina (Fisher Bioreagents, BP1760) o 40 µg/mL de kanamicina (Fisher Bioreagents, BP906-5), según fuese necesario, durante 16 horas a 37°C y 225 rpm. La extracción y purificación del ADN se llevó a cabo utilizando el kit GenElute Plasmid Miniprep (Sigma-Aldrich, PLN350), siguiendo las instrucciones del fabricante.

En las situaciones en las que se necesitó una mayor cantidad de ADN, las colonias individuales crecidas en placas de agar LB con antibiótico se cultivaron en 100 mL de medio líquido LB suplementado con 100 µg/mL de ampicilina o 40 µg/mL de kanamicina, durante 16 horas a 37°C y 225 rpm. El ADN fue extraído y purificado empleando el kit PureLink HiPure Plasmid Filter Midiprep (Invitrogen, K210015), siguiendo las instrucciones del fabricante.

La calidad y cantidad del ADN plasmídico resultante se evaluó mediante un espectrofotómetro Nanodrop 2000 (Thermo Scientific, ND-2000).

5.5 TRANSFECCIÓN DE CÉLULAS DE MAMÍFERO

Para llevar a cabo la sobreexpresión de las formas WT y mutantes catalíticas de la proteína DNMT2 humana mediante el sistema PiggyBac (ver sección 5.4.4), se sembraron inicialmente 150.000 células KO para *Dnmt2* por pocillo en placas de 6 pocillos (Corning, 3516), previamente tratadas con gelatina al 0,1%, utilizando medio pluripotente SL. Después de un periodo de 24 horas, se realizó la transfección simultánea de 1 µg de cada uno de los vectores PB junto con 1 µg de otro vector que contenía la transposasa (pBASE) (Wang et al., 2008) para facilitar su integración genómica (proporcionado por Jose C. Silva, de la Universidad de Cambridge, RU). Este proceso se llevó a cabo mediante el kit jetPRIME (Polyplus, 114-15) según las indicaciones del fabricante. Al día siguiente, se inició la selección con 250 µg/mL del antibiótico higromicina (Corning, 30-240-CR), manteniendo la presión selectiva hasta obtener ESCs con las formas sobreexpresadas de manera estable. La selección se prolongó hasta que un control con ESCs no transfectadas mostró un 100% de muerte celular.

El constructo diseñado para la generación de ESCs KO para *Dnmt2* mediante la técnica CRISPR/Cas9 (ver sección 5.4.7) fue introducido en las células pluripotentes CJ7 mediante el kit jetPRIME (Polyplus, 114-15), siguiendo las indicaciones del fabricante. Inicialmente, se sembraron 150.000 células por pocillo en placas de 6 pocillos (Corning, 3516), previamente tratadas con gelatina al 0,1%, en presencia de medio de cultivo pluripotente SL. Después de 24 horas en cultivo, se llevó a cabo la transfección con 2 µg del ADN plasmídico SpCas9-2A-Puro V2.0 que contenía la guía de ARN seleccionada ubicada en el exón 1 del gen *Dnmt2*. Las células transfectadas fueron seleccionadas de manera transitoria mediante el uso de 2 µg/mL de puomicina (Fisher Bioreagents, BP2956) durante dos días.

Para llevar a cabo el marcaje endógeno de la proteína DNMT2 en su extremo C-terminal, se realizó la transfección simultánea del vector que contenía la guía de ARN y el vector donador del marcaje 3xFLAG (ver apartado 5.4.8) en ESCs CJ7. Este procedimiento se

llevó a cabo utilizando el kit jetPRIME, siguiendo las instrucciones del fabricante. Inicialmente, se sembraron 110.000 células por pocillo en placas de 12 pocillos (Corning, 3513), previamente tratadas con gelatina al 0,1%, y cultivadas en presencia de medio de cultivo pluripotente SL. Después de 24 horas en cultivo, se procedió a la transfección simultánea con 540 ng del vector donador pFETCh-3xFlag-P2A-NeoR y 270 ng del vector que contenía la guía de ARN y la Cas9. Al día siguiente, se seleccionaron las células transfectadas de manera transitoria durante dos días con 2 µg/mL de puromicina y 400 µg/mL de G418 (Corning, 30-234-CR).

5.6 GENERACIÓN DE ESCs KO Y KI MEDIANTE CRISPR-Cas9

La línea de ESCs KO para *Dnmt2* se generó utilizando la herramienta de edición génica CRISPR/Cas9, como se detalló anteriormente (Ran et al., 2013). Para esto, se transfectaron las ESCs CJ7 con un vector que contenía un ARN guía, la enzima Cas9 y un gen de resistencia a puromicina, el cual se generó de acuerdo con las instrucciones de la sección 5.4.7, siguiendo el protocolo de transfección mencionado en la sección anterior. Después de la transfección y dos días de selección con puromicina (2 µg/mL), se procedió a la individualización de las células y a su siembra en densidad clonal, permitiendo así el aislamiento manual de colonias individuales originadas a partir de una única célula.

Para determinar la presencia de clones KO para *Dnmt2*, en primer lugar, se extrajo ADN genómico de los clones aislados (consultar la sección 5.8) para analizar la presencia de mutaciones por reacción en cadena de la polimerasa (PCR, por sus siglas en inglés) (ver apartado 5.9). Una vez seleccionados los clones candidatos, se amplificó la secuencia del gen de interés a partir del ADN genómico, que fue clonado en un vector pJET1.2. Después, las mutaciones generadas se detectaron mediante secuenciación de Sanger, empleándose el programa Sequencher (v5.4) para su análisis. La simulación de la secuencia de la proteína alterada se realizó con el programa SerialCloner (v2.6.1). Posteriormente, se extrajo ARN total (ver sección 5.10) de los clones seleccionados, con el fin de confirmar la reducción de los niveles de expresión de *Dnmt2* mediante la reacción en cadena de la polimerasa cuantitativa con transcripción inversa (RT-qPCR) (consultar la sección 5.12). Finalmente, se extrajo proteína total (ver apartado 5.13) para validar la ausencia de la proteína DNMT2 mediante *western blot* (WB) (consultar la sección 5.14).

La generación de la línea *Dnmt2*-3xFLAG KI también se llevó a cabo mediante la herramienta CRISPR/Cas9. Se transfectaron (ver sección 5.5) las ESCs CJ7 con el vector donador del fragmento 3xFLAG, generado tal y como se detalla en el apartado 5.4.8, además de con otro plásmido que contenía una guía de ARN dirigida al codón de terminación de la traducción del gen *Dnmt2* y la enzima Cas9 (previamente generado y validado en el laboratorio). Después, se seleccionaron las células durante dos días con puromicina (2 µg/mL) y G418 (400 µg/mL) y se procedió a la individualización de las células y a su siembra en densidad clonal, lo que permitió el aislamiento manual de colonias individuales originadas a partir de una única célula.

Para determinar la presencia de clones *Dnmt2*-3xFLAG KI se extrajo proteína total de los clones aislados (ver sección 5.13) y se validó la presencia de la etiqueta FLAG mediante WB (ver sección 5.14). Luego, se extrajo ADN genómico de los clones positivos para FLAG (ver sección 5.8), con el fin de evaluar la presencia de clones homocigotos para el marcaje endógeno de *Dnmt2*, por PCR (ver sección 5.9).

5.7 PRODUCCIÓN DE VIRUS Y TRANSDUCCIÓN CELULAR

5.7.1 Producción de partículas lentivirales y retrovirales

La encapsulación de constructos en partículas lentivirales y retrovirales se llevó a cabo siguiendo protocolos previamente publicados (Covelo-Molares et al., 2022; Fuentes-Iglesias et al., 2022). En el caso de las partículas lentivirales, se generaron a partir de los constructos pLKO.1, FUW, FUW-Tet-On, o el vector OSKM empleado para inducir la reprogramación celular a iPSCs (pHAGE2-EF1aFull-hOct4-F2A-hKlf4-IRES-hSox2-P2A-hcMyc-W-loxP) (Sommer et al., 2010). Este proceso se realizó mediante la transfección simultánea de estos vectores con los vectores empaquetadores psPAX2 (Addgene, 12260) y pMD2.G (Addgene, 12259) en células HEK293T. Por otro lado, se generaron partículas retrovirales a partir de los constructos pMX, lo que se llevó a cabo mediante la transfección individual de estos vectores en células Plat-E.

En ambos casos, se sembraron 8.000.000 de células en placas de cultivo de 150 mm (Corning, 430599) en presencia de medio de cultivo apropiado para estos tipos celulares (consultar la sección 5.3). Después de 24 horas en cultivo, se llevó a cabo la transfección utilizando polietilenimina (PEI; Sigma-Aldrich; 408727). Para esto, se preparó una solución de PEI y NaCl, que consistía en 16 μ L de PEI (22 mg/mL) por cada 1,6 mL de NaCl 150 mM, y se incubó en hielo hasta su uso. Además, se preparó una solución de ADN que incluía 20 μ g del vector lentiviral o retroviral. En el caso de la preparación de partículas lentivirales, se añadieron 10 μ g del vector empaquetador psPAX2 y 10 μ g del vector empaquetador pMDG2 por cada 1,6 mL de NaCl 150 mM.

A continuación, la solución de PEI fue añadida gota a gota sobre la solución de ADN, homogeneizando la mezcla mediante vórtex e incubándose a temperatura ambiente durante 30 minutos. Simultáneamente, se refrescó el medio de las células, al cual después de los 30 minutos de incubación se le añadió gota a gota la solución de transfección. Al día siguiente, se cambió el medio y se recolectaron los sobrenadantes que contenían virus después de 48 y 72 horas de transfección. Finalmente, se concentraron los sobrenadantes virales mediante centrifugación (3.500 xg durante 30 min a 4°C), utilizando filtros Amicon Ultra-15 (Millipore, UFC903024). Para evitar ciclos repetidos de congelación y descongelación, los sobrenadantes concentrados se almacenaron en alícuotas a -80°C.

5.7.2 Transducción celular

Para la transducción de fibroblastos con partículas lentivirales o retrovirales, se sembraron 5.263 células por cm^2 en presencia de medio completo. Después de 24 horas en cultivo, el medio fue reemplazado por medio completo suplementado con 6 μ g/mL de polibreno (Sigma-Aldrich, H9268) y 10-20 μ L de sobrenadantes virales concentrados. Al día siguiente, se realizó un cambio de medio, neutralizando el medio que contenía las partículas virales con lejía.

El proceso de transducción de ESCs o células tumorales con partículas virales se llevó a cabo de manera similar. Se sembraron 100.000 células por pocillo en placas de 12 pocillos (Corning, 3513), en presencia de SL para las ESCs, o medio completo para las células tumorales (consultar la sección 5.3). Después de 24 horas en cultivo, se cambió el medio de cultivo agregando 8 μ g/mL de polibreno y 20-50 μ L de sobrenadantes virales concentrados. Al día siguiente, se realizó un cambio de medio, neutralizando el medio que contenía las partículas lentivirales con lejía.

5.8 EXTRACCIÓN Y PURIFICACIÓN DE ADN GENÓMICO

Para la extracción y purificación de ADN genómico, se utilizó Puregene Cell Kit (Qiagen, 158043), siguiendo las instrucciones del fabricante. Brevemente, se resuspendieron 10^4 - 10^6 células en 200 μ L de solución de lisis celular, incubando el lisado durante una hora a 37°C. Posteriormente, se añadieron 3 μ g de ribonucleasa A (Thermo Scientific, EN0531) y la mezcla se incubó nuevamente durante 1 hora a 37°C. Luego, las muestras se mantuvieron en hielo durante 1 minuto, seguido de la adición de 66,6 μ L de solución de precipitación de proteínas. Las muestras se homogeneizaron durante 20 segundos mediante agitación con vórtex y se centrifugaron durante 10 minutos a 14.000 xg a RT. El sobrenadante resultante se transfirió a un tubo al que previamente se le habían añadido 200 μ L de isopropanol (Fisher Bioreagents, BP2618) y se mezcló por inversión. A continuación, las muestras se centrifugaron durante 5 minutos a 14.000 xg a RT y se eliminó el sobrenadante. El ADN sedimentado se resuspendió en 266,6 μ L de etanol al 70% (Sigma-Aldrich, 1.08543), y después de una incubación a RT de 5 minutos, se eliminó el exceso de líquido, permitiendo que el ADN se secase al aire durante 5 minutos. Finalmente, las muestras se resuspendieron en 66 μ L de solución de hidratación del ADN y se incubaron durante 1 hora a 65°C. La calidad y cantidad del ADN se determinaron utilizando un espectrofotómetro Nanodrop 2000 (Thermo Scientific, ND-2000).

5.9 REACCIÓN EN CADENA DE LA POLIMERASA

Con el fin de generar varios de los vectores plasmídicos empleados en este trabajo se llevó a cabo la clonación por reacción en cadena de la polimerasa (PCR) (ver secciones 5.4.4, 5.4.5 y 5.4.8). Para la amplificación de los fragmentos a clonar se emplearon polimerasas de alta fidelidad siguiendo las instrucciones de los distintos fabricantes y cuyas referencias se encuentran indicadas en las secciones correspondientes. Las secuencias de oligonucleótidos empleadas se encuentran en las **Tablas suplementarias 6 y 7**.

También se llevaron a cabo estrategias de genotipado por PCR para la identificación de clones *Dnmt2* KO y *Dnmt2* KI de ESCs, para lo que se empleó la mezcla maestra Dream Taq Green PCR 2X (Thermo Scientific, K1081). La presencia de mutaciones en el gen *Dnmt2* se analizó mediante una estrategia de amplificación con 3 oligonucleótidos, uno de los cuáles hibridaba en la región diana del ARN guía empleado (ver sección 5.4.7). Por otro lado, para la distinción de clones homocigotos para la inserción de la etiqueta 3xFLAG aguas abajo del gen *Dnmt2*, se llevó a cabo una PCR con dos oligonucleótidos, uno que hibrida en el último exón del gen y otro que hibrida aguas abajo del codón de terminación, en la región 3'-UTR. Las secuencias de los oligonucleótidos empleados se encuentran en la **Tabla suplementaria 6** y las condiciones de PCR se indican a continuación:

<i>Dnmt2</i> KO		
nº de ciclos	Temperatura (°C)	Tiempo
1	95	2 minutos
30	95	30 segundos
	60	30 segundos
	72	1 minuto
1	72	10 minutos
-	4	-

<i>Dnmt2</i> KI		
nº de ciclos	Temperatura (°C)	Tiempo
1	95	2 minutos
30	95	30 segundos

	58.4	30 segundos
	72	1 minuto
1	72	10 minutos
-	4	-

Las reacciones de amplificación se realizaron en el termociclador SimpliAmp (Applied Biosystems, A24811).

5.10 EXTRACCIÓN Y PURIFICACIÓN DE ARN

Para los análisis de expresión génica mediante transcripción inversa seguida de reacción en cadena de la polimerasa cuantitativa (RT-qPCR) se realizó la extracción de ARN total utilizando E.Z.N.A Total RNA Purification Kit 1 (Omega Bio-Tek, R6834-02), siguiendo las instrucciones proporcionadas por el fabricante. La calidad y cantidad de ARN fueron evaluadas mediante un espectrofotómetro Nanodrop 2000 (Thermo Scientific, ND-2000).

Para la secuenciación masiva de ARN (RNA-seq) o en ensayos destinados a la detección de moléculas de ARNt, se purificó el ARN total utilizando el reactivo TRIzol (Invitrogen, 15596018), junto con tubos Phasemaker (Invitrogen, A33248), siguiendo las recomendaciones del fabricante. En resumen, se añadieron 0,3-0,6 mL de TRIzol directamente al cultivo celular por cada 10^5 - 10^7 células, respectivamente, y se homogeneizó la muestra por pipeteo. Luego, las muestras se transfirieron a los tubos Phasemaker, se añadieron 0,2 mL de cloroformo (Fisher Chemical, C/4920/15) por cada mL de TRIzol y la mezcla se homogeneizó por inversión. Tras incubarse a RT durante 2-3 minutos, se centrifugaron durante 5 minutos a 16.000 xg y 4°C. La fase acuosa se recogió y el ARN se precipitó añadiendo 0,5 mL de isopropanol (Fisher Bioreagents, BP2618) por cada mL de TRIzol. Después de mezclar por inversión, las muestras se incubaron durante 10 minutos a RT y se centrifugaron durante 10 minutos a 12.000 xg y 4°C. El ARN sedimentado se lavó con 1 mL de etanol 75% frío (Sigma-Aldrich, 1.08543), se centrifugó durante 5 minutos a 7.500 xg y 4°C, se descartó el sobrenadante y se dejó secar al aire durante 5 minutos a RT. Finalmente, el ARN se solubilizó en agua ultrapura (Invitrogen, 10977-035). La cantidad y calidad del ARN se analizaron utilizando un fluorómetro Qubit (Invitrogen, Q33226) junto con los kits Qubit HS Assay (Invitrogen, Q32855) y Qubit RNA IQ Assay (Invitrogen, Q33222), respectivamente.

Para su conservación a largo plazo, las muestras de ARN se almacenaron a -80°C hasta su uso.

5.11 OBTENCIÓN DE ADN COMPLEMENTARIO

El ADN complementario (ADNc) utilizado para el análisis de expresión génica por RT-qPCR se generó mediante una reacción en la que se combinó una cantidad específica de ARN total con la mezcla maestra qScript cDNA SuperMix (Quantabio, P/N 84034), siguiendo las recomendaciones del fabricante. La reacción se incubó durante 5 minutos a 25°C, 30 minutos a 42°C, 5 minutos a 85°C y se estabilizó a 4°C en un termociclador.

Para la obtención del ADNc específico para el ARNt-Asp utilizado en el análisis de los niveles de metilación de esta molécula, se empleó la transcriptasa inversa Maxima H Minus (Thermo Scientific, EP0753), siguiendo las instrucciones proporcionadas por el fabricante. En este caso, se utilizó un oligonucleótido específico cuya secuencia se proporciona en la **Tabla suplementaria 9**.

5.12 REACCIÓN EN CADENA DE LA POLIMERASA CUANTITATIVA CON TRANSCRIPCIÓN INVERSA

Con el fin de analizar la expresión génica mediante transcripción inversa seguida de reacción en cadena de la polimerasa cuantitativa (RT-qPCR), se siguió el protocolo que se describe a continuación. De manera individual y para cada una de las regiones de interés, se preparó una mezcla de ADNc (ver apartado 5.11), la mezcla maestra PowerUp SYBR Green (Applied Biosystems, A25741) y una combinación de oligonucleótidos específicos para la amplificación de cada región, ambos a una concentración de 1 μ M. Estas mezclas específicas para la amplificación de las regiones de interés se cargaron en placas de reacción de 384 pocillos (Axygen, PCR-384M2-C), las cuales fueron selladas con películas PlateMax UltraClear (Axygen, UC-500). La reacción se llevó a cabo en el sistema de PCR en tiempo real QuantStudio 5 (Applied Biosystems, A33621) utilizando las siguientes condiciones:

n° de ciclos	Temperatura (°C)	Tiempo
1	98	10 minutos
40	95	15 segundos
	55	15 segundos
	72	40 segundos
	95	5 segundos
1	65	
1	(subiendo 5°C/minuto hasta alcanzar 97°C)	1 minuto
1	97	15 segundos
-	4	mantener

Los datos obtenidos fueron analizados según el método $2^{-\Delta\Delta C_t}$ a partir de triplicados técnicos y los valores de expresión relativa se normalizaron utilizando genes control y la muestra de referencia correspondiente. Todas las figuras se generaron utilizando el programa GraphPad Prism 8. Las secuencias de los cebadores empleados se proporcionan en la **Tabla suplementaria 8**.

5.13 EXTRACCIÓN Y PURIFICACIÓN DE PROTEÍNA TOTAL

Los extractos de proteína total se obtuvieron utilizando el sistema de tampón de lisis RIPA (Santa Cruz Biotechnology, sc-24948), siguiendo las recomendaciones del fabricante. Brevemente, las células en cultivo fueron individualizadas, centrifugadas durante 5 minutos a 300 xg y resuspendidas en DPBS (Corning, 21-031-CV) para eliminar el medio de cultivo residual. Posteriormente, se recolectaron mediante centrifugación y se resuspendieron en tampón de lisis RIPA, al cual se le añadieron los inhibidores de proteasas (PMSF y PIC) y de fosfatasas (ortovanadato de sodio) del kit. Las muestras se incubaron en hielo durante 30 minutos, agitándolas cada 2-3 minutos con un agitador vórtex. Finalmente, se centrifugaron las muestras durante 10 minutos a 10.000 xg y 4°C y el sobrenadante obtenido se almacenó a -80°C hasta su uso.

5.14 WESTERN BLOT

Los niveles totales de proteínas específicas fueron determinados mediante la técnica de *western blot* (WB), según se describe a continuación. En primer lugar, se cuantificó la cantidad de proteína presente en cada muestra utilizando el kit de ensayo de proteínas BCA Pierce (Thermo Scientific, 23227), siguiendo las instrucciones del fabricante. Posteriormente, se combinó la cantidad deseada de proteína con el tampón de carga LaemmLi 1X [8% de SDS (Fisher Bioreagents, BP1311), 40% de glicerol, 5% de β -mercaptoetanol (Sigma-Aldrich,

M6250), 0,04% de azul de bromofenol (Sigma-Aldrich, B8026) y 240 mM de Tris-HCl pH 6,8 (Bio Basic, SD8122)] en agua ultrapura filtrada a través de un filtro de 0,22 μ m (Fisherbrand, 15181489) mediante una jeringuilla (Terumo, 0197)], procediéndose a la desnaturalización durante 5 minutos a 95°C. Las muestras se cargaron en geles Novex WedgeWell 4–20% tris-Glycine Gel (Invitrogen, XP04205BOX o XP04200BOX) o en geles de acrilamida/bisacrilamida (Sigma-Aldrich, A3574) preparados en el laboratorio, para su resolución mediante electroforesis. A continuación, las proteínas fueron transferidas a membranas de fluoruro de polivinilideno (PVDF) (Cytiva, 10600023), durante 16 horas a 4°C. Al día siguiente, las membranas se lavaron con agua ultrapura y se tiñeron con una solución compuesta por reactivo de Ponceau S (0,1%) (Sigma-Aldrich, P3504) y ácido acético glacial (1%) (Fisher Scientific, 10171460), durante 3 minutos a temperatura ambiente y en agitación, lo que permitió la visualización de los patrones de proteína total. Posteriormente, se realizaron tres lavados consecutivos con agua ultrapura, durante 5 minutos a temperatura ambiente y en agitación, seguidos de otros tres lavados de 5 minutos a temperatura ambiente y en agitación, con TBS-T [tampón salino Tris 1X (Fisher Bioreagent, BP2471-1) suplementado con 0,1% de TWEEN 20 (Sigma-Aldrich, P1379)]. Con el fin de evitar uniones inespecíficas, las membranas fueron bloqueadas durante 1 hora a RT y en agitación, con leche (LabScientific, M0841) diluida al 5% en TBS-T, después de lo cual se lavaron tres veces con TBS-T, como se describió previamente. Seguidamente, se incubaron las membranas con anticuerpos primarios específicos para las proteínas de interés, diluidos en leche al 0,5% en TBS-T. El exceso de anticuerpo primario se eliminó mediante 3 lavados de 15 minutos a temperatura ambiente y en agitación con TBS-T. Posteriormente, las membranas se incubaron durante 45 minutos a 4°C y en agitación con los correspondientes anticuerpos secundarios conjugados con peroxidasa de rábano (HRP), permitiendo su detección mediante quimioluminiscencia. Para eliminar el exceso de anticuerpo secundario, las membranas se lavaron 5 veces durante 10 minutos en TBS-T y se incubaron con el sustrato para inmunotransferencia ECL Pierce (Thermo Scientific, 32106). La señal de quimioluminiscencia fue detectada mediante películas de rayos X Fuji (Fujifilm, 100NIF) y el procesador de películas Fuji FPM-100A. En los casos en los que se realizaron cuantificaciones de los niveles de proteína relativos al control, se empleó el software Image J (v1.33). Los anticuerpos empleados se proporcionan en la **Tabla suplementaria 12**.

5.15 ENSAYOS DE REPROGRAMACIÓN CELULAR

5.15.1 Reprogramación celular de fibroblastos con OSKM

Para llevar a cabo los ensayos de reprogramación celular con el objetivo de generar células troncales pluripotentes inducidas (iPSCs), se empleó la expresión forzada de los factores de transcripción OSKM (OCT4, SOX2, KLF4 y cMYC) en MEFs *Oct4*-GFP, que contienen un reportero GFP bajo el control del promotor del marcador de pluripotencia *Oct4*, o en fibroblastos adultos. Inicialmente, los fibroblastos fueron sembrados en medio de cultivo para fibroblastos, a una confluencia de 10.000 células por pocillo en placas de 24 pocillos (Corning, 3524), previamente tratadas con gelatina al 0,1%. Después de 24 horas, los fibroblastos fueron transducidos con shRNAs en el caso de experimentos de pérdida de función y/o con constructos de sobreexpresión en experimentos de ganancia o rescate de función. Simultáneamente, se llevó a cabo la transducción con partículas lentivirales que contenían los factores de reprogramación OSKM. Al día siguiente, se cambió el medio de cultivo de fibroblastos por medio pluripotente (SL), que fue renovado diariamente hasta la finalización del experimento.

Para evaluar la eficiencia de reprogramación, se llevaron a cabo diversas aproximaciones que incluyeron: el recuento de colonias pluripotentes en campo claro (BF), el conteo de colonias positivas para la fosfatasa alcalina después de la tinción (consultar la sección 5.18.1), el recuento de colonias positivas para *Oct4*-GFP y la determinación del número de células individuales positivas para *Oct4*-GFP mediante citometría de flujo (ver la sección 5.21).

5.15.2 Reprogramación de pre-iPSCs

Para llevar a cabo la reprogramación de intermediarios epiteliales avanzados (pre-iPSCs) a iPSCs, las células pre-iPSCs fueron generadas según se describió anteriormente (Costa et al., 2013). Brevemente, se transdujeron fibroblastos embrionarios de ratón (MEFs) que contenían un reportero GFP controlado por el promotor del marcador de pluripotencia *Nanog* (*Nanog*-GFP) con partículas retrovirales que contenían los factores de reprogramación OSKM. Después de tres días, se cambió el medio a medio pluripotente (SL) y, tras 12 días en estas condiciones, se seleccionó la población GFP negativa, que se correspondía con las pre-iPSCs, mediante un clasificador de células activado por fluorescencia (FACS). Posteriormente, la población *Nanog*-GFP negativa se transfectó con un vector PB vacío o que codificaba NANOG, utilizando Lipofectamina 3000, y las células que albergaban el vector fueron seleccionadas con higromicina durante 12 días.

Después de completar la selección, las pre-iPSCs que expresaban un vector vacío o la proteína NANOG fueron transducidas con shRNAs contra *Dnmt2* o *Luciferasa* (control). Después de 24 horas, para asegurar la expresión estable de los shRNAs, se inició una selección con puomicina (2 µg/mL) (Fisher Bioreagents, BP2956) que se mantuvo hasta que las células pre-iPSCs no transducidas mostraron un 100% de muerte. Luego, se sembraron 20.000 células por pocillo en placas de 12 pocillos con medio SL. Al día siguiente, se refrescó el medio cambiándolo por el medio de reprogramación 2i+LIF, basado en medio N2B27 [50% de medio DMEM/F-12 (Sigma-Aldrich, D6421) y 50% de medio Neurobasal (Gibco, 21103049), suplementado con 1X de B-27 (Gibco, 17504044), 1X de N-2 (Gibco, 17502048), 2 mM de L-Glutamina (Gibco, 25030-081), 110 U/mL de penicilina/estreptomicina (Gibco, 15070-063) y 9,125 ng/mL de β-mercaptoetanol) (Sigma-Aldrich, M6250)]; suplementado con 1.000 U/mL de LIF (producido y probado en el laboratorio), 1 µM del inhibidor de la quinasa MEK (PD0325901) (Sigma-Aldrich, PZ0162) y 3 µM del inhibidor de la quinasa GSK3β (CHIR99021) (Sigma-Aldrich, SML1046). Finalmente, después de 10 días de reprogramación, la eficiencia de reprogramación se evaluó mediante el conteo de colonias positivas para GFP.

5.15.3 Reprogramación celular directa a iHSCs

La reprogramación de MEFs y HDFs hacia células troncales hematopoyéticas inducidas (iHSCs) se llevó a cabo siguiendo protocolos previamente publicados (Daniel et al., 2016; Silverio-Alves et al., 2019). En primer lugar, se realizó la transducción de los fibroblastos con shRNAs contra el gen *Dnmt2* o *Luciferasa* como control, para los experimentos de pérdida de función; y con constructos pFUW-Tet-On, cuya expresión es activada en respuesta a doxiciclina (DOX) en presencia del transactivador M2rtTA, en el caso de los experimentos de ganancia de función. Al día siguiente, las células transducidas con shRNAs fueron seleccionadas con 2 µg/mL de puomicina (Fisher Bioreagents, BP2956) para garantizar la expresión estable de los constructos, mientras que en las células transducidas con constructos FUW-Tet-On se indujo la expresión de estos mediante la adición de 1 µg/mL de DOX (Sigma-Aldrich, D9891) al medio. La selección con puomicina se mantuvo hasta que un control celular no transducido con

shRNAs mostró un 100% de muerte y la inducción con DOX se mantuvo durante todo el experimento.

Posteriormente, se sembraron 20.000 células por pocillo en placas de 12 pocillos previamente tratadas con gelatina al 0,1%. Después de 48 horas en cultivo en medio completo, se procedió a la transducción de MEFs o HDFs con partículas lentivirales que albergaban constructos de expresión constitutiva (pFUW), codificantes para los factores de transcripción GATA2, GFI1B y FOS (GGF), proporcionados por el Dr. Carlos-Filipe Pereira, de la Universidad de Lund (Suecia). Después de dos días, el medio de cultivo de los fibroblastos fue reemplazado por un medio de reprogramación hematopoyética compuesto por Myelocult M5300 (STEMCELL Technologies, 05350) en el caso de MEFs o Myelocult H5100 (STEMCELL Technologies, 05150) en el caso de los HDFs, ambos suplementados con la hormona hidrocortisona (STEMCELL Technologies, 74142). Tras 15 días en condiciones de reprogramación, con cambios de medio cada 2 días, se procedió a evaluar la eficiencia del proceso mediante citometría de flujo (sección 5.21) o conteo de colonias.

5.16 ENSAYOS DE DIFERENCIACIÓN

5.16.1 Diferenciación a cuerpos embrioides

Las ESCs se cultivaron en suspensión en medio SL sin LIF para fomentar la formación de agregados celulares esféricos en tres dimensiones conocidos como cuerpos embrioides (EBs), capaces de recapitular en cultivo algunas características del desarrollo embrionario *in vivo* (Doetschman et al., 1985). Para esto, las ESCs WT o KO para *Dnmt2*, o transducidas con *shDnmt2* o *shLuci*, se sembraron a una densidad de 155.000 células/mL en placas Petri de baja adherencia (Deltalab, 200200). Estas se mantuvieron en cultivo durante 10 días, refrescando el medio cada 2 días. Para evitar perturbar los agregados ya generados, los EBs se recolectaron y precipitaron por gravedad, permitiendo la aspiración del sobrenadante y el cambio del medio. La evaluación de las diferencias morfológicas entre las condiciones estudiadas se llevó a cabo capturando imágenes representativas en los días 6, 8 y 10 del ensayo, utilizando un microscopio invertido Olympus IX51 con una cámara Olympus DP72. Además, en el día 10 se extrajo ARN (ver sección 5.10), a partir del cual se generó ADNc (ver sección 5.11) y se analizaron los niveles de expresión de marcadores de diferenciación temprana por RT-qPCR (ver sección 5.12).

5.16.2 Generación de gastruloides

Las ESCs fueron diferenciadas en cultivo formando agregados celulares tridimensionales conocidos como gastruloides. Estas estructuras imitan algunos de los procesos observados durante la gastrulación *in vivo*, incluyendo cambios en la elongación, polarización y expresión génica (Beccari et al., 2018). Para esto, se siguió un protocolo previamente publicado (Baillie-Johnson et al., 2015). De manera resumida, se sembraron inicialmente 400 ESCs WT o KO para *Dnmt2* por pocillo en placas de 96 pocillos con fondo en U (Corning, 3788). Estas células fueron cultivadas durante 120 horas en medio N2B27 (composición detallada en la sección 5.15.2), con un tratamiento de CHIR99021 a 3 μ M aplicado entre las 48 y 72 horas. El medio se cambió diariamente, con cuidado de no perturbar los agregados formados. Después de 120 horas, se tomaron imágenes representativas y se extrajo ARN (ver sección 5.10), a partir del cual se generó ADNc (ver sección 5.11) para analizar los cambios transcripcionales en ausencia de DNMT2 por RT-qPCR (ver sección 5.12). Las imágenes se adquirieron utilizando un microscopio invertido Olympus IX51 con una cámara Olympus DP72

y se emplearon para cuantificar el tamaño de las estructuras mediante el software Image J (v1.33). Para calcular la eficiencia en la generación de estructuras gastruloides, se comparó el porcentaje de agregados que alcanzaron una morfología similar, así como su tamaño (n=8 o 4 grupos de 8 agregados cada uno).

5.16.3 Diferenciación a células mesenquimales

Con el fin de evaluar los defectos en el cambio de identidad celular durante el proceso de diferenciación en ausencia de DNMT2, se llevó a cabo un ensayo de diferenciación hacia células del linaje mesenquimal. En este caso, se sembraron 20.000 ESCs WT o KO para *Dnmt2* por pocillo en placas de 12 pocillos tratadas previamente con gelatina al 0,1%. Tras 24 horas en condiciones estándar de cultivo para células pluripotentes, estas fueron tratadas con 5 ng/ μ L de TGF- β (Sigma-Aldrich, H8541) de manera consecutiva durante 5 días, tanto en presencia como ausencia de LIF (SL $\pm$ LIF). Las diferencias morfológicas entre condiciones fueron evaluadas mediante la toma de imágenes a días 0 y 5 del ensayo, con un microscopio invertido Olympus IX51 con una cámara Olympus DP72. Además, se examinó la capacidad de diferenciación hacia el linaje mesenquimal de estas células mediante el análisis de los niveles totales de la proteína α -SMA por WB con un anticuerpo específico contra esta proteína (ver sección 5.14 y **Tabla suplementaria 12**), a días 0 y 5 del ensayo, en presencia y ausencia de LIF.

5.16.4 Diferenciación a fibroblastos

Para confirmar los defectos observados durante el proceso de diferenciación bajo la ausencia de DNMT2, se llevó a cabo la generación de células con características similares a fibroblastos a partir de ESCs WT y KO para *Dnmt2*, adaptando un protocolo previamente publicado (Ruiz et al., 2012). Brevemente, se sembraron 620.000 ESCs de cada condición en placas Petri de baja adherencia (Deltalab, 200200). Estas, fueron mantenidas durante cuatro días en cultivo en presencia de medio SL sin LIF, de manera similar a como se hizo durante la generación de EBs (ver sección 5.16.1). Posteriormente, los agregados generados se transfirieron a placas de 100 mm de diámetro (Corning, 430167), previamente tratadas con gelatina al 0,1%, para favorecer su adhesión, también en presencia de medio SL sin LIF. Después de dos días, las células adheridas fueron individualizadas y sembradas en placas no tratadas con gelatina en presencia de medio de enriquecimiento para fibroblastos (DMEM/F-12, suplementado con 10% de suero fetal bovino, 0,5 mM de L-Glutamina, 0,1 mM de aminoácidos no esenciales y 55 μ M de β -mercaptoetanol). De esta manera, se favoreció el enriquecimiento de la población celular en células tipo fibroblasto. El ensayo se mantuvo bajo estas condiciones hasta el día 18, momento en el que las poblaciones celulares WT y KO para *Dnmt2* fueron teñidas con cristal violeta (ver sección 5.18.2) para evaluar su viabilidad. Finalmente, las diferencias entre condiciones fueron evaluadas mediante la toma de imágenes en días clave del experimento, utilizadas para el análisis morfológico. Las imágenes se tomaron con un microscopio invertido Olympus IX51 equipado con una cámara Olympus DP72.

5.16.5 Diferenciación a mesendodermo y neuroectodermo

Con el fin de analizar la función de la actividad catalítica de DNMT2 durante la adquisición de nuevas identidades celulares, se llevó a cabo un ensayo de diferenciación de ESCs a los linajes de neuroectodermo (NE) y mesendodermo (ME), como ya se describió previamente (Thomson et al., 2011). En este caso, se utilizaron ESCs WT y KO para *Dnmt2*, así como dos líneas celulares *Dnmt2* KO sobre las que se había inducido la expresión estable

de las versiones humanas catalíticamente activa (WT) e inactiva (C79A) mediante transfección con el sistema PiggyBac (ver sección 5.5). Brevemente, se sembraron 57.000 ESCs por pocillo, por triplicado, en placas de 12 pocillos previamente tratadas con gelatina al 0,1%, en presencia de medio N2B27 (composición detallada en la sección 5.15.2). Durante las primeras 48 horas este medio fue suplementado con 10^3 U/mL de LIF, 10 ng/mL de BMP4 (Gibco, PHC9534) y PD0325901 1 μ M (Sigma-Aldrich, PZ0162). Posteriormente, las células se mantuvieron durante 48 horas más en medio N2B27 sin suplementos. Finalmente, se procedió a su diferenciación mediante la suplementación del medio N2B27 con CHIR99021 3 μ M o con 500 nM de ácido retinoico (RA; Santa Cruz Biotechnology, sc-200898), hacia ME y NE, respectivamente. Se procedió a la extracción de ARN (ver apartado 5.10) en los días 0 y 5 del ensayo, a partir del cual se obtuvo ADNc (ver apartado 5.11), que fue empleado para evaluar las diferencias entre condiciones durante la diferenciación a estos linajes celulares, mediante RT-qPCR (ver apartado 5.12).

5.17 ENSAYOS DE FORMACIÓN DE COLONIAS

5.17.1 Ensayo de formación de colonias pluripotentes

Con el objetivo de evaluar el impacto de la pérdida de función de DNMT2 en la capacidad de autorrenovación de las células pluripotentes, se llevaron a cabo ensayos de formación de colonias siguiendo el protocolo descrito a continuación. Se sembraron 800 ESCs por pocillo, ya fueran transducidas con shRNAs contra *Dnmt2* o *Luciferasa*, o bien células WT y KO para *Dnmt2*, sobre gelatina al 0,1% (Sigma-Aldrich, G9391), en medio pluripotente (SL). Al día siguiente, se realizó un cambio de medio por SL o medio con suero sin factor de crecimiento LIF. Se añadió medio fresco diariamente y después de 4 días en cultivo, se llevó a cabo una tinción para la fosfatasa alcalina (ver apartado 5.18.1). Posteriormente, se contaron las colonias y se clasificaron según su tinción en: colonias no diferenciadas (morfología compacta y tinción homogénea), colonias parcialmente diferenciadas (morfología irregular y tinción heterogénea) o colonias diferenciadas (morfología plana y pérdida completa de tinción).

5.17.2 Ensayo de formación de colonias en agar blando

Para analizar los cambios en la capacidad de crecimiento independiente de anclaje de las líneas humanas MCF-7 y MDA-MB-231 de cáncer de mama tras la pérdida de función de DNMT2, se llevó a cabo un ensayo de formación de colonias sobre agar blando con células transducidas con shRNAs contra *DNMT2* o *Luciferasa*, como control. Es importante destacar que este ensayo se realizó en placas de 6 pocillos y que, en el momento de la siembra, todas las soluciones se mantuvieron a una temperatura constante de 42°C. Previamente al momento de la siembra, se prepararon dos soluciones de agar (Sigma-Aldrich, A5431) en agua ultrapura, una al 1% y otra al 0,6%, que fueron esterilizadas y mantenidas a 4°C hasta el inicio del experimento.

En primer lugar, se disolvió la solución de agar al 1% en un microondas y se mantuvo a 42°C durante 30 minutos en un baño de agua. A continuación, se preparó una mezcla 1:1 de agar al 1% y DMEM 2X, compuesto por: 2,68g de DMEM en polvo (Gibco, 52100-021), 0,74g de NaHCO₃ (Thermo Scientific, 10128190), 20% de FBS (Gibco, 10270106) previamente inactivado a 56°C durante 30 minutos y 200 U/mL de penicilina/estreptomicina (Gibco, 15070-063). Se vertieron 1,5 mL de esta mezcla en cada pocillo, dejándola secar durante 30 minutos a RT. Posteriormente, se disolvió la solución de agar al 0,6% en un microondas y se mantuvo, al igual que en el caso anterior, a 42°C durante 30 minutos en un baño de agua. Durante este

tiempo de incubación, se procedió al conteo de células para su dilución a una densidad de 20.000 células por mL en DMEM 2X.

Finalmente, se preparó una mezcla 1:1 de agar al 0,6% y células diluidas en DMEM 2X, vertiéndose 1,5 mL de esta mezcla, que se dejaron secar al menos 30 minutos a RT. Posteriormente, se añadieron 100 μ L de medio completo en la superficie de la capa de agar al 0,6% y células. Durante el transcurso del experimento, se realizó un cambio de medio de cultivo cada 3 días, añadiendo como máximo 100 μ L de medio completo cada vez. Tras 15 días, se llevó a cabo el conteo de las colonias formadas.

5.18 TINCIONES

5.18.1 Tinción contra la fosfatasa alcalina

Las células pluripotentes se caracterizan por presentar elevados niveles de la enzima fosfatasa alcalina (AP), siendo este marcador ampliamente utilizado en el campo de pluripotencia (Stefkova et al., 2015). Para evaluar la pluripotencia después de los ensayos de reprogramación celular a iPSCs y para detectar posibles defectos en el mantenimiento de la pluripotencia tras la pérdida de la función de *Dnmt2*, se determinó la presencia de AP utilizando el kit Leucocyte Alkaline Phosphatase (Sigma-Aldrich, 86R-1KT), siguiendo las instrucciones del fabricante.

De manera resumida, las células en cultivo se lavaron con DPBS con Ca^{+2} y Mg^{+2} (Sigma-Aldrich, D8662) y se fijaron durante 30 segundos a RT con una solución de fijación. Luego, se lavaron con agua desionizada y se incubaron con una solución de tinción recién preparada durante 15 minutos a RT en la oscuridad. Finalmente, las células se lavaron nuevamente con agua desionizada durante 2 minutos y se dejaron secar al aire. Las imágenes se adquirieron utilizando un microscopio invertido Olympus IX51 con una cámara Olympus DP72.

Composición de las soluciones de fijación y tinción (los volúmenes indicados corresponden a soluciones para una placa de 6 pocillos):

Solución de fijación	
Citrato (Sigma-Aldrich, 86R-1KT)	2,5 mL
Acetona (Fisher Chemical, A/0560/17)	6,5 mL
Formaldehído 37% (Sigma-Aldrich, F8775)	0,8 mL
Solución de tinción	
Solución de nitrito de sodio (Sigma-Aldrich, 86R-1KT)	0,2 mL
Solución FRV-alcalina (Sigma-Aldrich, 86R-1KT)	0,2 mL
Solución alcalina de naftol AS-BI (Sigma-Aldrich, 86R-1KT)	0,2 mL
Agua desionizada	9 mL

5.18.2 Tinción con cristal violeta

La tinción con cristal violeta realizada para evaluar la viabilidad celular durante la diferenciación de ESCs hacia fibroblastos (ver sección 5.16.4) se realizó tal y como se detalla a continuación. En primer lugar, se retiró el medio de las células para proceder a su lavado con DPBS. Estas fueron fijadas con paraformaldehído (Sigma-Aldrich, P6148) diluido al 4% en DPBS con CaCl_2 y MgCl_2 (Sigma-Aldrich, D8662), durante 15 minutos en la oscuridad. Después, se realizaron dos lavados con DPBS y se procedió a la tinción con cristal violeta (Sigma-Aldrich, C0775-25G) diluido al 0,05% en agua desionizada, durante 15 minutos a RT. Finalmente, el exceso de tinción fue eliminado mediante lavados repetidos con DPBS y la tinción se dejó secar al aire durante 16 horas.

5.19 INMUNOFLUORESCENCIA

Con el fin de analizar la abundancia y localización de las proteínas OCT4 y SOX2 en ESCs, se siguió el protocolo de inmunofluorescencia detallado a continuación. En primer lugar, se sembraron 10.000 ESCs por pocillo, las cuales fueron previamente transducidas con shRNAs contra *Dnmt2* o *Luciferasa* como control, en una placa de 48 pocillos (Corning, 3548), utilizando medio de cultivo SL. Al día siguiente, las células fueron fijadas con 300 μL de paraformaldehído (Sigma-Aldrich, P6148) diluido al 4% en DPBS, durante 15 minutos a temperatura ambiente y en oscuridad.

Posteriormente, las células fueron lavadas tres veces durante 5 minutos en DPBS a RT. Seguidamente, se procedió a la permeabilización y bloqueo en un solo paso mediante la incubación durante 30 minutos a 37°C en 300 μL de un tampón compuesto por 5% de suero fetal bovino (FBS; Gibco, 10270106) y 0,3% de Tritón X-100 (Sigma-Aldrich, T8787) en DPBS. Luego, se realizó un nuevo lavado con DPBS y se incubaron las células durante 16 horas a 4°C con los anticuerpos primarios correspondientes diluidos 1:1000 en un tampón compuesto por 1% de albúmina de suero bovino (BSA; Fisher Bioreagents, BP1605) y 0,3% de Tritón X-100 en DPBS.

Tras esta incubación, se llevaron a cabo tres lavados de 5 minutos con DPBS, seguidos de una incubación de 2 horas a temperatura ambiente, en la oscuridad, con los anticuerpos secundarios conjugados con fluoróforos, diluidos 1:500 en el mismo tampón. Además, se añadió 1 μM de 4',6-diamidino-2-fenilindol (DAPI; Sigma-Aldrich, D9542) para la tinción nuclear. Finalmente, las células fueron lavadas tres veces durante 5 minutos en DPBS y conservadas a 4°C y en oscuridad.

Las imágenes de inmunofluorescencia fueron adquiridas utilizando el microscopio invertido Leica DMI6000 B, con una magnificación de 20X, y ajustes constantes para cada fluoróforo en las distintas condiciones experimentales. Posteriormente, las imágenes fueron procesadas empleando el software Leica LAS X (v3.4.2.18368). Los detalles de los anticuerpos y las condiciones empleadas se encuentran en la **Tabla suplementaria 12**.

5.20 ANÁLISIS DE CICLO CELULAR

Para analizar posibles diferencias en el ciclo celular tras el silenciamiento o eliminación de *Dnmt2* en ESCs, se llevó a cabo una tinción con yoduro de propidio, siguiendo el siguiente protocolo. Inicialmente, se sembraron 250.000 células por pocillo en duplicado, para todas las condiciones, en placas de 6 pocillos (Corning, 3516), sobre gelatina al 0,1%, utilizando medio de cultivo estándar para células pluripotentes (SL).

Al día siguiente, las células fueron recogidas y centrifugadas durante 5 minutos a 500 xg, seguido de un lavado con DPBS. Posteriormente, las células se fijaron con etanol (Sigma-Aldrich, 1.08543) frío diluido al 70% en agua y se incubaron durante 30 minutos a 4°C. El etanol fue añadido gota a gota mientras se agitaban las muestras a 16 rpm con la ayuda de un mezclador vórtex (Fisherbrand, 11746744). Tras la fijación, las células se atemperaron durante 5 minutos a temperatura ambiente, se resuspendieron mediante inversión y se centrifugaron de nuevo durante 5 minutos a 850 xg. El sedimento celular fue lavado dos veces con DPBS y centrifugado durante 5 minutos a 850 xg.

Después, las células sedimentadas se resuspendieron en solución de tinción, compuesta por yoduro de propidio (50 µg/mL) (Molecular Probes, P3566) y 10 µg/mL de ribonucleasa A (Thermo Scientific, EN0531) en DPBS. La suspensión se incubó durante 30 minutos a 37°C en oscuridad. Finalmente, la intensidad del yoduro de propidio se analizó mediante citometría de flujo (ver siguiente sección).

5.21 CITOMETRÍA DE FLUJO

Para analizar mediante citometría de flujo la intensidad de fluorescencia del reportero GFP en ensayos de reprogramación, o del yoduro de propidio en los análisis de ciclo celular, se llevó a cabo el siguiente procedimiento. Inicialmente, las células fueron recogidas y centrifugadas durante 5 minutos a 300 xg. Luego, se lavaron con DPBS y se mantuvieron en hielo en tubos de citometría (Corning, 352058) hasta su análisis. La intensidad de fluorescencia se cuantificó utilizando un citómetro FACSCalibur (BD Biosciences) o un citómetro LSR Fortessa (BD Biosciences). Para esto, se analizaron un total de 10.000 eventos por condición. Los datos generados fueron analizados mediante el software FlowJo (v7.6.1).

En cuanto al análisis de los niveles del marcador hematopoyético de membrana CD49f después de la inhibición de *DNMT2* durante la reprogramación celular de HDFs a iHSCs, se siguió el siguiente procedimiento. Inicialmente, se levantaron las células y se añadió 1 mL de DPBS suplementado con un 2% de FBS, por muestra. Las células fueron recogidas mediante centrifugación a 350 xg. Luego, el sobrenadante fue descartado por inversión y al remanente se le añadieron 0,2 µg de anticuerpo α -CD49f (Invitrogen, 25-0495-82) a cada muestra. Esta mezcla fue incubada durante 30 minutos a temperatura ambiente en oscuridad, y posteriormente centrifugada a 350 xg. Nuevamente, se descartó el sobrenadante por inversión y, a continuación, la intensidad de fluorescencia se cuantificó mediante el empleo de un citómetro LSR Fortessa (BD Biosciences). Los detalles sobre el anticuerpo empleado para la tinción del marcador CD49f, así como sus condiciones de uso, se proporcionan en la **Tabla suplementaria 12**.

5.22 ENSAYOS DE PROLIFERACIÓN

Para el análisis de los patrones de proliferación de ESCs y líneas tumorales tras el silenciamiento o eliminación de *DNMT2* en cultivo, se llevó a cabo el siguiente procedimiento. Se sembraron 100.000 células por pocillo y por triplicado para cada condición, sobre placas de 12 pocillos (Corning, 3513), previamente tratadas con gelatina diluida al 0,1% (Sigma-Aldrich, G9391) en el caso de las ESCs.

Después de 2 días en cultivo, se recogieron las células y se sembró un porcentaje en nuevos pocillos para continuar el ensayo. El resto se empleó para contar el número total de células en una cámara de Neubauer (Hirschmann, 8100103). Este proceso se repitió cada 2 días hasta la finalización del experimento, calculando el número total de células acumuladas a lo largo del ensayo.

Con el objetivo de analizar si el silenciamiento de *Dnmt2* afecta la proliferación de los fibroblastos en cultivo, ya sea en presencia o ausencia de OSKM, se llevaron a cabo análisis de proliferación siguiendo procedimientos similares a los descritos anteriormente. En ambos casos, se sembraron en duplicado 10.000 fibroblastos por pocillo en placas de 24 pocillos (Corning, 3524), en condiciones de medio somático o pluripotente. El recuento del número total de células se realizó a lo largo del ensayo utilizando una cámara de Neubauer.

5.23 ENSAYO DE CIERRE DE HERIDAS

La capacidad de migración de las células tumorales MCF-7 y MDA-MB-231 tras la inhibición de *DNMT2* fue evaluada mediante un ensayo de cierre de heridas, conocido como *wound healing* (WH). Inicialmente, se sembraron 300.000 células en placas de 12 pocillos (Corning, 3513), con cuatro pocillos por condición, utilizando medio completo. Después de 24 horas en cultivo, alcanzando el 100% de confluencia celular en el pocillo, se creó una hendidura vertical o herida en la monocapa de células. Posteriormente, las células fueron trasladadas a la cámara de incubación del microscopio Celldiscoverer 7 (ZEISS), manteniéndose a 37°C con un 5% de CO₂ durante la duración del experimento. Se evaluaron las diferencias en la tasa de migración celular mediante la toma de imágenes y el cálculo del área relativa de la herida en diferentes puntos temporales a lo largo del ensayo. Las mediciones de las áreas se llevaron a cabo utilizando el software Image J (v1.33).

5.24 DETERMINACIÓN DE m⁵C EN ARN

5.24.1 Análisis de la abundancia de m⁵C por ELISA

Para cuantificar los niveles globales de la marca 5-metilcitosina (m⁵C) en el ARN de ESCs WT o KO para *Dnmt2*, se llevó a cabo un ensayo de inmunoabsorción ligado a enzima (ELISA). Se utilizó el kit MethylFlash 5-mC RNA Methylation ELISA Easy (Epigentek, P-9009-48), siguiendo las instrucciones del fabricante. Para la detección de la señal fluorométrica se midió absorbancia a 530 nm empleando un lector de placas Mithras LB 940 (Berthold Technologies).

5.24.2 Análisis de la abundancia de m⁵C por HPLC-MS/MS

Para investigar la potencial función catalítica de DNMT2 en las fracciones de ARNm y ARN largo no codificante (ARNlnc), se determinaron los niveles de m⁵C en ESCs WT y KO para *Dnmt2* mediante cromatografía líquida de alta resolución acoplada a la espectrometría de masas (HPLC-MS/MS). En este proceso, se sembraron por triplicado 500.000 ESCs WT y KO para *Dnmt2* por pocillo en placas de 6 pocillos, previamente tratadas con gelatina al 0,1%, en presencia de medio pluripotente SL suplementado con 1 µM del inhibidor de la quinasa MEK (PD0325901) (Sigma-Aldrich, PZ0162) y 3 µM del inhibidor de la quinasa GSK3β (CHIR99021) (Sigma-Aldrich, SML1046). Después de 48 horas en cultivo, se recolectaron las muestras con 600 µL de TRIzol por pocillo. La extracción y purificación de ARN total se realizaron según lo indicado en la sección 5.10. Posteriormente, se empleó el kit Turbo DNA-free (Invitrogen, AM1907), para eliminar posibles restos de ADN, siguiendo las instrucciones del fabricante. El ARN purificado se cuantificó con un espectrofotómetro Nanodrop 2000 y se evaluó su integridad por electroforesis en gel de agarosa al 1%.

Posteriormente, se utilizó el kit RNA clean and concentrator-5 (Zymo Research, R1015), siguiendo las instrucciones del fabricante para obtener dos fracciones de ARN a partir

de la muestra inicial: una de fragmentos menores de 200 nucleótidos y otra de fragmentos mayores de 200 nucleótidos. En esta última fracción, se llevó a cabo el análisis de los niveles de m5C. La eliminación del ARNr de la fracción de interés se realizó utilizando el kit RiboMinus Eukaryote Probe Mix v2 (Ambion, A15020), siguiendo las recomendaciones del fabricante.

Las muestras luego se enviaron a las instalaciones externas del Centro de Regulación Genómica (CRG) en Barcelona para el análisis HPLC-MS/MS. En este proceso, las muestras fueron digeridas con la mezcla maestra Nucleoside Digestion Mix (New England BioLabs, M0649S) y los ribonucleósidos se purificaron con HyperSep Hypercarb SPE Spin Tips (Thermo Fisher Scientific, 60109-209). El análisis HPLC-MS/MS se llevó a cabo utilizando un espectrómetro de masas Orbitrap Eclipse Tribrid (Thermo Fisher Scientific) acoplado a un EASY-nLC 1200 (Thermo Fisher Scientific).

5.24.3 Análisis de metilación por secuenciación Sanger

Para analizar los niveles de metilación en las citosinas de las moléculas de ARNt en ESCs WT y KO para *Dnmt2* o durante el proceso de reprogramación celular somática, se realizó la extracción de ARN total con TRIzol (ver sección 5.10) y posteriormente se sometió a un tratamiento con bisulfito. Este procedimiento induce la desaminación de las citosinas no metiladas, convirtiéndolas en uracilos que, tras la síntesis de ADNc, se detectan como timinas en procesos de secuenciación. En contraste, las citosinas metiladas no experimentan desaminación como resultado de este tratamiento.

Para llevar a cabo la conversión con bisulfito, se utilizó 1 µg de ARN total por muestra con el kit EZ RNA Methylation (Zymo Research, R5001), siguiendo las instrucciones del fabricante. Después de evaluar la calidad y cantidad del ARN mediante un espectrofotómetro Nanodrop 2000 (Thermo Scientific, ND-2000), se procedió a la síntesis de ADNc específico de ARNt-Asp mediante la adición de un cebador específico a la reacción (ver apartado 5.11). Posteriormente, se amplificó el ADNc de ARNt-Asp por PCR utilizando cebadores específicos (detallados en **Tabla suplementaria 9**) y se visualizó el producto de amplificación en un gel de agarosa al 2% mediante electroforesis.

El ADNc de ARNt-Asp convertido con bisulfito se purificó del gel utilizando el kit GenElute Gel Extraction (Sigma-Aldrich, NA1111) y se clonó en un vector pJET1.2 con el kit CloneJet PCR (Thermo Scientific, K1231), siguiendo las recomendaciones del fabricante en ambos casos. Posteriormente, se llevó a cabo la transformación de bacterias competentes con el vector generado, permitiendo su crecimiento en presencia de antibiótico durante 16 horas a 37°C y 225 rpm. Al día siguiente, se purificaron los vectores utilizando el kit GenElute Plasmid Miniprep (Sigma-Aldrich, PLN350), siguiendo las recomendaciones del fabricante. Luego, se validó el grado de metilación de las posiciones de interés mediante secuenciación de Sanger y se analizaron los resultados con el software Sequencher (v5.4).

Los oligonucleótidos empleados para la secuenciación se detallan en la **Tabla suplementaria 11**.

5.25 SECUENCIACIÓN MASIVA DE ARN Y ANÁLISIS

El ARN total destinado a la secuenciación masiva (RNA-seq) se extrajo de acuerdo con el procedimiento detallado en la sección 5.10. Los perfiles transcripcionales de MEFs transducidos con shRNAs contra *Dnmt2* o *Luciferasa* a día 5 de reprogramación celular

mediante OSKM, así como los de ESCs WT, KO para *Dnmt2* y líneas de rescate con las versiones mutantes de DNMT2 (KO+C79A y KO+R371H), fueron analizados mediante esta técnica. La empresa Novogene, encargada de la secuenciación, evaluó la calidad del ARN utilizando el sistema Agilent Bioanalyzer 2100. Luego, se realizó un protocolo para eliminar el ARNr utilizando el kit RiboZero (Illumina, 20040526) antes de la preparación de las librerías de secuenciación. Estas se generaron de manera no específica para cada cadena en el caso de las ESCs (*non-stranded*) y específica para cada cadena (*stranded*) en el caso de las muestras de reprogramación OSKM a día 5 a partir de fibroblastos. La secuenciación del ARN se llevó a cabo utilizando las plataformas Illumina HiSeq4000 para las muestras de ESCs e Illumina NovaSeq6000 para las muestras de reprogramación, generando lecturas emparejadas de 150 pares de bases en ambos casos.

El análisis bioinformático de los resultados de secuenciación fue realizado por mí, siguiendo protocolos establecidos en el campo. Inicialmente, se evaluó la calidad de los datos de secuenciación mediante el programa FastQC (v0.11.5), utilizando la escala de Phred como referencia. Posteriormente, se realizó el alineamiento de las lecturas de alta calidad contra el genoma de ratón (Ensembl GRCm39, versión 104) utilizando el programa STAR (v2.7.6). Las lecturas alineadas se asignaron a los transcritos correspondientes mediante el programa htseq-count (v0.13.5), especificando el parámetro de superposición "*intersection-nonempty*" para evitar asignar una lectura a más de un gen. El análisis diferencial de la expresión génica se llevó a cabo con el paquete DESeq2 (v1.32.0) en R (v4.2.2), utilizando *shLuci* como muestras de control en el caso de reprogramación celular a día 5, y las *Dnmt2* WT como control en el análisis de perfiles transcripcionales de ESCs. Las matrices de datos se anotaron con el paquete BiomaRT (v2.52.0) en R, agregando atributos útiles para futuros análisis (nombre oficial de los genes, cromosoma, posición de inicio y fin, hebra y tipo de transcrito).

Es importante señalar que el análisis de elementos transponibles (TEs), se realizó mediante el programa SQuIRE (Yang et al., 2019b), que facilitó los pasos de alineamiento, asignación y análisis diferencial.

Los datos de secuenciación masiva de ARN tratado con bisulfito en ESCs WT y KO para *Dnmt2*, publicados por Legrand y colaboradores (Legrand et al., 2017), fueron analizados en R, empleando los paquetes diseñados por sus autores y siguiendo el flujo de trabajo establecido por el grupo de origen. En este caso, variamos el criterio FDR, de 0,05 a 0,01.

Los datos de RNA-seq previamente publicados por otros grupos y utilizados en esta tesis doctoral fueron analizados mediante el mismo flujo de trabajo descrito anteriormente. Los datos crudos están disponibles en la base de datos de acceso libre NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>), bajo los códigos de acceso proporcionados en la **Tabla suplementaria 13**.

5.26 ANÁLISIS *IN SILICO* DE LOS DATOS TRANSCRIPTÓMICOS

5.26.1 Análisis de componentes principales

Para comparar las diferencias entre los perfiles transcripcionales en distintas condiciones celulares a partir de los datos de RNA-seq, se llevó a cabo un análisis de componentes principales (PCA). Inicialmente, los datos transcripcionales normalizados (*normalized counts* o lecturas normalizadas) se importaron al programa Cluster 3.0. Luego, se restó la media de expresión génica a cada condición mediante la opción *center genes*, asegurando que el valor medio de cada fila sea 0. Posteriormente, se normalizaron los datos

utilizando la opción *normalize genes* y se aplicó el análisis de componentes principales a matrices mediante la opción *apply PCA to arrays*. Finalmente, los resultados se proyectaron en una representación tridimensional utilizando el programa Origin (Origin Pro 2017b v9.4.0.220).

5.26.2 Matrices de correlación y agrupaciones jerárquicas

La representación de los datos de expresión en matrices de correlación con agrupaciones jerárquicas mediante mapas de calor se llevó a cabo empleando el programa Morpheus (<https://software.broadinstitute.org/morpheus/>). Los datos de expresión asociados a cada gen para cada condición fueron importados al programa para luego establecer el grado de similitud de las muestras en base al coeficiente de correlación de Pearson.

5.26.3 Análisis de ontología génica y vías KEGG

Para determinar los procesos y vías afectados tras la pérdida de función de DNMT2 durante el proceso de generación de iPSCs o en ESCs, se realizaron análisis de ontología génica (GO) y análisis de vías de señalización KEGG (Kyoto Encyclopedia of Genes and Genomes). En ambos casos, se empleó la herramienta de anotación funcional DAVID (Database for Annotation, Visualization, and Integrated Discovery; <https://david.ncifcrf.gov/>) (Sherman et al., 2022). También se llevaron a cabo análisis de enriquecimiento de procesos biológicos empleando el sistema de clasificación PANTHER (<https://www.pantherdb.org/>) (Mi et al., 2019; Thomas et al., 2022). En estas herramientas se importaron distintas listas de genes desregulados a nivel transcripcional tras la pérdida de función de DNMT2 o de su actividad catalítica, que fueron analizadas contra una lista de referencia que albergaba todos los genes expresados en *Mus musculus*. Los códigos de los términos GO y KEGG empleados a lo largo de este trabajo se proporcionan en las **Tablas suplementarias 14 y 15**, respectivamente.

5.26.4 GSEA

Para analizar las diferencias en el enriquecimiento de los perfiles transcripcionales de las muestras de estudio con respecto a un conjunto de genes dado, se utilizó el programa GSEA (Gene Set Enrichment Analysis; v4.3.2). Los datos de expresión obtenidos por RNA-seq se importaron en formato de texto delimitado por tabulaciones (.txt) al programa. Luego, se compararon los perfiles transcripcionales de las distintas condiciones (por ejemplo, WT vs KO) con respecto a listas de genes definidas, que también se importaron al programa. A continuación, se establecieron una serie de parámetros, según se indica en la siguiente tabla:

Parámetro	Valor
Número de permutaciones	1000
Tipo de permutación	<i>gene_set</i>
Estadístico de enriquecimiento	<i>weighted</i>
Métrica de clasificación de genes	<i>Ratio_of_classes</i>

Tras la ejecución del análisis, el programa generó los gráficos de enriquecimiento, y calculó tanto el valor de enriquecimiento normalizado (NES), como la tasa de descubrimientos falsos (FDR).

5.27 IDENTIFICACIÓN DEL INTERACTOMA DE DNMT2 EN ESCs

5.27.1 Inmunoprecipitación y nLC-MS/MS

Con el propósito de identificar las proteínas que interactúan con DNMT2 en ESCs, se llevó a cabo una inmunoprecipitación de proteína total en ESCs KI para *Dnmt2*, generadas con marcaje endógeno homocigótico 3xFLAG aguas abajo del gen *Dnmt2* (ver sección 5.6), y en ESCs WT, como control. Posteriormente, se realizó un análisis de espectrometría de masas en tándem por cromatografía de nanolíquidos (nLC-MS/MS). A continuación, se describe detalladamente el procedimiento.

Las ESCs fueron cultivadas hasta obtener 7 placas de 150 mm de diámetro (Corning, 430599) con una confluencia del 80-90%. Posteriormente, se llevó a cabo su individualización, seguida de una centrifugación (5 minutos a 300 xg) y un lavado con DPBS frío. Después de recoger las células nuevamente por centrifugación, se procedió al lisado mediante la adición de 10,5 mL de un tampón de lisis con inhibidores de proteasas añadidos en fresco (la composición detallada del tampón se proporciona al final de esta sección).

Las células fueron incubadas en este tampón durante 10 minutos a 4°C y posteriormente sometidas a sonicación utilizando el sonicador Sonifier de Branson, aplicando 5 ciclos (1 segundo ON y 10 segundos OFF) con una amplitud del 30%. Luego, se incubaron durante 10 minutos en hielo, seguido de una centrifugación durante 10 minutos a 10.000 xg y 4°C. Los sobrenadantes fueron recolectados y dializados en 75 volúmenes de tampón de diálisis en presencia de inhibidores de proteasas (la composición detallada del tampón se proporciona al final de esta sección) durante 16 horas a 4°C, con agitación constante, utilizando casetes de diálisis de Slide-A-Lyzer G2, 10K MWCO (Thermo Scientific, 87731).

Posteriormente, se recogieron los dializados y se centrifugaron durante 20 minutos a 18.000 xg y 4°C. El sobrenadante conteniendo los lisados de proteína total fue almacenado a -80°C y la proteína fue cuantificada utilizando el kit de ensayo de proteínas BCA Pierce (Thermo Scientific, 23227), siguiendo las instrucciones del fabricante. Con el objetivo de eliminar posibles uniones inespecíficas, se realizó un pre-aclarado de la muestra añadiendo 150 µg de bolas magnéticas (Dynabeads Protein G; Invitrogen, 1004D) por cada mg de proteína. La mezcla se incubó durante 1 hora y 30 minutos a 4°C en un rotador (Grant-bio, PTR-35). Posteriormente, las bolas magnéticas fueron retiradas utilizando el imán DynaMag-2 (Invitrogen, 12321D), y las muestras se almacenaron a -80°C hasta su uso.

La inmunoprecipitación (IP) se llevó a cabo utilizando extractos de proteína total de ESCs WT y KI. Para ello, se utilizaron 30 mg de proteína por reacción de IP, separándose el 10% de cada mezcla como input. Luego, se añadieron 12 µg de anticuerpo primario α -FLAG (Sigma-Aldrich, F1804) en cada IP, incubándose las mezclas durante 16 horas a 4°C en rotación constante. Al día siguiente, se prepararon 10 µL de bolas magnéticas por µg de anticuerpo (stock a 30 µg/µL), en tampón de IP en presencia de inhibidores de proteasas (la composición detallada del tampón se proporciona al final de esta sección), y se realizaron 2 lavados con el mismo tampón y con la ayuda del imán DynaMag-2. Después, se resuspendieron las bolas en 200 µL de tampón de IP por reacción y se añadieron a los complejos proteína y anticuerpo previamente formados. La nueva mezcla se incubó durante 3 horas a 4°C, en rotación constante. Utilizando el imán DynaMag-2, se recogieron los complejos de bolas magnéticas+anticuerpo+proteína y se lavaron un total de 3 veces en 8 mL de tampón de IP con inhibidores de proteasas recién añadidos durante 10 minutos a 4°C, bajo agitación constante.

Tras los lavados, los complejos se resuspendieron en 800 μ L de tampón de IP y se incubaron durante otros 10 minutos a 4°C en agitación. Finalmente, se recogieron con el imán DynaMag-2, se resuspendieron en 200 μ L de tampón de carga LaemmLi 2X para su elución (la composición detallada del tampón se proporciona al final de esta sección), y se incubaron durante 5 minutos a 95°C. Luego, se recogieron las bolas empleando el imán DynaMag-2, se guardó el sobrenadante y se repitió el proceso añadiendo otros 200 μ L de tampón de carga LaemmLi 2X a las bolas. Se guardó de nuevo el sobrenadante, que se mezcló con el anterior, y se realizó un ciclo final de elución con 100 μ L de tampón LaemmLi 1X. Este tercer sobrenadante también se mezcló con el anterior, obteniendo un total de 500 μ L de muestra, que se conservaron a -80°C hasta su uso.

Finalmente, se procedió a concentrar las muestras utilizando filtros Amicon Ultra de 0,5 mL (Sigma-Aldrich, UFC500308), siguiendo las instrucciones del fabricante. Se obtuvieron aproximadamente 50 μ L de muestra, que se resolvieron en una electroforesis a 100 voltios en un gel de Tris-Glicina con un gradiente de 20-4% (Novex WedgeWell; Invitrogen, XP04200BOX). Posteriormente, el gel se selló y se envió a las instalaciones externas en CiC bioGUNE para su análisis mediante nLC-MS/MS.

- Tampón de lisis: 25 mM de Tris-HCl pH 7,5 (Fisher Bioreagents, BP1757), 20% de glicerol (Sigma-Aldrich, G5516), 100 mM de KCl (Fisher Chemical, P/4240/60), 0,5 mM de EDTA (Fisher BioReagents, BP2482) y 10% Igepal (Sigma-Aldrich, I3021) en agua ultrapura (Invitrogen, 10977-035), filtrado a través de un filtro de 0,22 μ m (Fisherbrand, 15181489) empleando una jeringuilla (Terumo, 0197). Se añadieron los siguientes inhibidores de proteasas en fresco: DTT (1 mM) (Sigma-Aldrich, 43816), PMSF (0,2 mM) (Sigma-Aldrich, 93482) y el cóctel inhibidor de proteasas (PIC) (Sigma-Aldrich, P8340) (dilución 1:1000).

- Tampón de diálisis: 20 mM de HEPES pH 7,9 (Fisher Bioreagents, BP299), 25% de glicerol, 1,5 mM de MgCl₂ (Sigma-Aldrich, M1028) y 0,2 mM de EDTA en agua ultrapura filtrado. Se añadieron los siguientes inhibidores de proteasas en fresco: DTT (1 mM) (Sigma-Aldrich, 43816), PMSF (0,2 mM) (Sigma-Aldrich, 93482) y el cóctel inhibidor de proteasas (PIC) (Sigma-Aldrich, P8340) (dilución 1:1000).

- Tampón de IP: 20 mM de HEPES pH 7,9, 25% de glicerol, 1,5 mM de MgCl₂, 0,2 mM de EDTA y 0,02% de Igepal en agua ultrapura filtrado. Se añadieron los siguientes inhibidores de proteasas en fresco: DTT (1 mM) (Sigma-Aldrich, 43816), PMSF (0,2 mM) (Sigma-Aldrich, 93482) y el cóctel inhibidor de proteasas (PIC) (Sigma-Aldrich, P8340) (dilución 1:1000).

- Tampón de carga LaemmLi 2X: 4% de SDS (Fisher Bioreagents, BP1311), 20% de glicerol, 10% de β -mercaptoetanol (Sigma-Aldrich, M6250), 0,02% de azul de bromofenol (Sigma-Aldrich, B8026) y 120 mM de Tris-HCl pH 6,8 (Bio Basic, SD8122) en agua ultrapura filtrado.

5.27.2 Análisis de redes de interacción

Para visualizar y analizar la relación entre las proteínas que interactúan con DNMT2 en ESCs, se utilizó el programa Cytoscape (v3.8.2) (Shannon et al., 2003). Para identificar posibles interacciones entre proteínas, se empleó la aplicación STRING (Szklarczyk et al., 2021) y los datos obtenidos se integraron en Cytoscape. Además, se incorporaron los datos de expresión (FC y FDR) de las proteínas que interactúan con DNMT2, los cuales fueron obtenidos por

RNA-seq bajo su ausencia. En la representación visual, el color de los círculos indica el valor de cambio (FC), mientras que el tamaño indica el valor de la significancia estadística ($-\text{Log}_2$ del FDR).

El análisis de los fenotipos asociados a la alteración de las proteínas que interactúan con DNMT2 se llevó a cabo utilizando la herramienta en línea Enrichr (<https://maayanlab.cloud/Enrichr/>).

5.28 OBTENCIÓN Y ANÁLISIS DE DATOS DE PACIENTES CON CÁNCER

Los datos del Atlas del Genoma del Cáncer (TCGA) relativos a las mutaciones de DNMT2 en pacientes con distintos tipos de cáncer fueron obtenidos de la base de datos GDC Data Portal (<https://portal.gdc.cancer.gov/>). Posteriormente, fueron analizados en GraphPad Prism 8.

Para los datos de supervivencia de pacientes con cáncer en relación con la expresión de DNMT2, se utilizaron los datos depositados por el consorcio Pan-Cancer, los cuales fueron obtenidos del repositorio del programa Kaplan-Meier Plotter (<https://kmplot.com/analysis/>) (Nagy et al., 2021).

La información sobre la expresión de DNMT2 en tejidos sanos y tumorales se obtuvo de las bases de datos GTEx (<https://gtexportal.org/home/>) y OncoDB (Tang et al., 2022). Por otro lado, los datos relativos a la expresión de DNMT2 en líneas celulares tumorales se adquirieron del portal Depmap (<https://depmap.org/portal/home/#/>).

5.29 ANÁLISIS FILOGENÉTICOS

Con el fin de comparar las secuencias de las proteínas NSUN y DNMT de *Mus musculus* y *Homo sapiens* llevamos a cabo el alineamiento de las mismas, obtenidas en la base de datos Uniprot (<https://www.uniprot.org/>), mediante el programa Clustal Omega (v1.2.4) (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Después, construimos un árbol filogenético mediante el método de unión de vecinos en la propia base de datos Uniprot, que fue visualizado sin raíz mediante la herramienta en línea iTOL (Letunic y Bork, 2021).

En el caso de la comparativa de las proteínas DNMT en las secuencias de organismos vertebrados, las secuencias fueron obtenidas de la base de datos del NCBI. Posteriormente se llevó a cabo el alineamiento de estas secuencias mediante el programa MEGA X (Kumar et al., 2018), aplicando el algoritmo ClustalW. Posteriormente se construyó un árbol filogenético mediante el método de máxima verosimilitud, que fue visualizado mediante la herramienta en línea iTOL (Letunic y Bork, 2021).

5.30 ANÁLISIS ESTADÍSTICOS

La mayor parte de análisis estadísticos de esta tesis doctoral se llevaron a cabo mediante el programa GraphPad Prism 8. Para evaluar la significancia estadística de las diferencias entre grupos, se utilizaron distintas pruebas estadísticas según el caso.

La distribución de los conjuntos de datos se evaluó mediante las pruebas de normalidad de Shapiro-Wilk (si el número de datos por condición era menor de 50) o la prueba de Kolmogorov-Smirnov (si el número de valores era superior a 50).

En los análisis de expresión o eficiencia de reprogramación, se emplearon pruebas ANOVA de un factor o Kruskal-Wallis no pareadas, en función de la distribución de datos, si el número de condiciones era mayor de dos; y pruebas t de Student no pareadas cuando dos condiciones eran analizadas. En caso de no conocer el signo de la diferencia, se emplearon pruebas t de Student de dos colas, mientras que en casos en los que sí se conocía se emplearon pruebas de una cola (por ejemplo, en el análisis de expresión de un gen tras su inhibición). Con el fin de determinar las diferencias específicas entre grupos, las pruebas ANOVA, fueron complementadas con pruebas de comparaciones múltiples (por ejemplo, Dunnet, Sidak, etc.).

Para los casos en los que se analizaban distintas variables además de distintas condiciones (por ejemplo, proliferación celular, formación de colonias, entre otras), se emplearon pruebas ANOVA de dos factores no pareadas, si el número de grupos por análisis era mayor de dos; y múltiples pruebas t de Student, con corrección FDR para comparaciones múltiples, si el número de grupos a analizar era igual a dos.

Para analizar las diferencias en la supervivencia de pacientes con cáncer se realizaron pruebas Kaplan-Meier, empleándose una prueba Mantel-Cox para el cálculo del valor *P* y del cociente de riesgo (HZ).

Las pruebas estadísticas empleadas en esta tesis doctoral se encuentran indicadas en la leyenda de cada una de las figuras. Las diferencias fueron consideradas estadísticamente significativas cuando la probabilidad de que se debiesen al azar era menor del 5%, es decir, cuando el valor *P* era menor de 0,05.

5.31 RECURSOS BIBLIOGRÁFICOS

La revisión bibliográfica fue elaborada empleando las plataformas digitales de libre acceso Google Académico (<https://scholar.google.es/>) y Pubmed (<https://www.ncbi.nlm.nih.gov/pubmed/>), para la consulta de las referencias utilizadas. La introducción de estas referencias fue llevada a cabo mediante el software EndNote X9 (<https://endnote.com/>).

5.32 PROGRAMAS Y ALGORITMOS

Programa	Fuente	Enlace
CRISPOR	(Concordet y Haeussler, 2018)	http://crispor.tefor.net
BiomaRT v2.52.0	(Durinck et al., 2009)	https://bioconductor.org/packages/release/bioc/html/biomaRt.html
ChemDraw (v22.2)	PerkinElmer	https://www.perkinelmer.com/es/product/chemoffice-chemoffice
Cluster v3.0	Stanford University	http://bonsai.hgc.jp/~mdehoon/software/cluster/
Cytoscape v3.8.2	(Shannon et al., 2003)	https://cytoscape.org
DAVID	Lab of Human Retrovirology and Immunoinformatic	https://david.ncicrf.gov/tools.jsp
DESeq2 v1.32.0	(Love et al., 2014)	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Enrichr	Ma'ayan Lab	https://maayanlab.cloud/Enrichr/
FastQC v0.11.5	Babraham Institute	https://www.bioinformatics.babraham.ac.uk/projects/fastqc
FlowJo v7.6.1	FlowJo, LLC	https://www.flowjo.com/
GraphPad Prism 8	GraphPad	https://www.graphpad.com/
GSEA v4.2.2	Broad Institute	https://www.gsea-msigdb.org/gsea
htseq-count v0.13.5	(Putri et al., 2022)	https://htseq.readthedocs.io/en/master/

Image J v1.33	NIH	https://imagej.nih.gov/ij
Inkscape v1.3	Free Software Foundation, Inc.	https://inkscape.org/
iTOL v6.8.1	(Letunic y Bork, 2021)	https://itol.embl.de/
Kaplan-Meier Plotter	Semmelweis Univ. (Bioinformatics Department)	https://kmplot.com/analysis/
Leica LAS X v3.4.2.18368	Leica Microsystems	https://www.leica-microsystems.com
MEGA X	(Kumar et al., 2018)	https://www.megasoftware.net/
Morpheus	Broad Institute	https://software.broadinstitute.org/morpheus/
Origin Pro-2017 v9.4.0.220	OriginLab	https://www.originlab.com/Origin
PANTHER	National Science Foundation	https://www.pantherdb.org/
R v4.2.2	The R Foundation	https://www.r-project.org
Sequencher 5.4	Gene Codes Corporation	http://www.genecodes.com/
SerialCloner v2.6.1	Serial Cloner	http://serialbasics.free.fr/Serial_Cloner.html
SQulRE	(Yang et al., 2019b)	https://github.com/wyang17/SQulRE
STAR v2.7.6	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
STRING	(Szklarczyk et al., 2021)	https://string-db.org/
EndNote X9	The EndNote Team	https://endnote.com/

6. RESULTADOS

6.1 ANÁLISIS DEL PAPEL DE DNMT2 EN LA REPROGRAMACIÓN CELULAR SOMÁTICA

6.1.1 La presencia de DNMT2 es necesaria para la generación de iPSCs

Con el propósito de investigar la función de DNMT2 en la reprogramación celular somática (SCR) hacia células pluripotentes, llevamos a cabo una estrategia de pérdida de función mediante el empleo de ARN de interferencia.

Inicialmente, se transdujeron fibroblastos embrionarios de ratón (MEFs) con dos shRNAs específicos dirigidos contra el gen *Dnmt2*, los cuales fueron previamente generados y validados en células troncales embrionarias (ESCs) de ratón (consultar la sección 6.2). También se incluyó un shRNA dirigido contra el gen *Luciferasa* como control. Estos MEFs contienen la proteína fluorescente verde (GFP) bajo el control del promotor del gen *Oct4*, que se expresa exclusivamente en células pluripotentes (Lengner et al., 2007). Simultáneamente, realizamos la transducción con los factores de reprogramación OCT4, SOX2, KLF4 y cMYC, abreviados en conjunto como OSKM (Takahashi y Yamanaka, 2006) (Figura R1a).

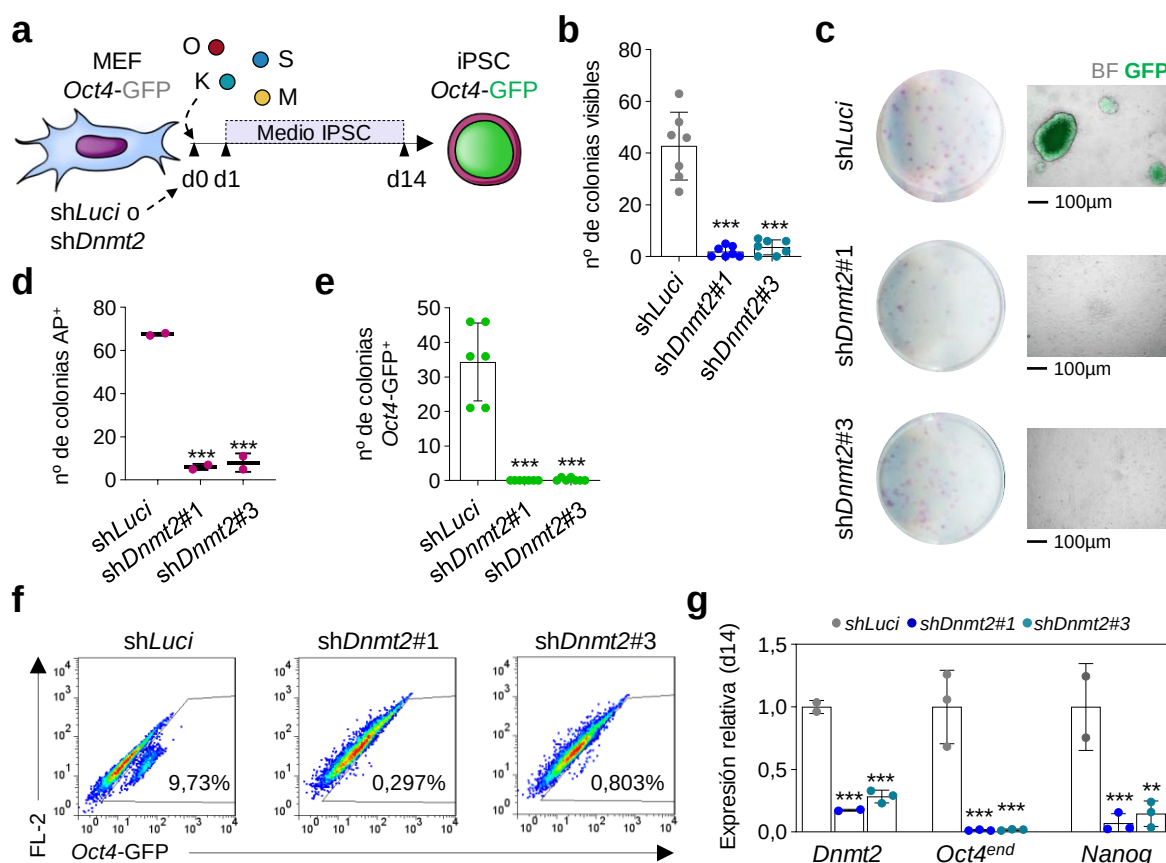


Figura R1. El silenciamiento de *Dnmt2* compromete la generación de iPSCs a partir de MEFs. a) Representación esquemática de la estrategia de reprogramación celular utilizada para la generación de iPSCs. La reprogramación de MEFs *Oct4-GFP* se indujo mediante la transducción de los factores OSKM junto con shRNAs dirigidos contra *Dnmt2* o *Luciferasa* (control). Se evaluaron las colonias pluripotentes (BF, AP+, GFP+) después de 14 días en medio de reprogramación (medio IPSC). b) Recuento de colonias con morfología pluripotente en campo claro (BF). c) Imágenes representativas de tinción de fosfatasa alcalina (AP, izquierda) y de BF combinada con fluorescencia (BF/GFP, derecha). Escala: 100 µm. d) Cuantificación del número de colonias AP+. e) Número de colonias GFP+. f) Cuantificación del número de células GFP+ mediante citometría de flujo para las condiciones indicadas a día 14 de SCR. g) Análisis de los niveles de expresión de *Dnmt2* y de los marcadores de pluripotencia *Oct4* (forma endógena, para diferenciarla de la sobreexpresada en el cóctel de reprogramación OSKM) y *Nanog* en las condiciones indicadas a día 14 de SCR. Los niveles de expresión se muestran relativos a los niveles de *U6*

y al control (shLuci). En b, d, e y g, los datos se representan como media $\pm$ desviación estándar y el número de replicados biológicos se indica por puntos. Se utilizó una prueba ANOVA de un factor no pareada con corrección de Dunnet para comparaciones múltiples para calcular las diferencias estadísticamente significativas, indicadas como **: $p < 0,01$ y ***: $p < 0,001$.

Después de 14 días en presencia de medio de reprogramación, evaluamos el efecto de la supresión de *Dnmt2* en la generación de células pluripotentes inducidas (iPSCs) (**Figura R1a**). Para esto, cuantificamos el número total de colonias con morfología pluripotente en imágenes de campo claro (BF), así como el número de colonias positivas para los marcadores de pluripotencia fosfatasa alcalina (AP) y *Oct4*-GFP. De manera destacada, observamos una reducción significativa en el número de colonias con morfología de iPSCs en presencia de shRNAs dirigidos contra *Dnmt2* en comparación con el grupo de control (shLuci) (**Figura R1b**). Esta reducción se reflejó tanto en el número de colonias AP+ como en el número de colonias GFP+ en comparación con shLuci (**Figuras R1c-R1e**).

Seguidamente, examinamos el efecto de la supresión de *Dnmt2* en la expresión de *Oct4* a nivel de célula individual mediante citometría de flujo. En consonancia con los datos anteriores, se observó que la reducción de *Dnmt2* conllevaba una disminución significativa de los niveles del marcador GFP (0,297% y 0,803%), en comparación con las células transducidas con el shRNA dirigido contra el gen *Luciferasa* (9,73%) (**Figura R1f**). Después, se determinaron los niveles endógenos de expresión de los marcadores de pluripotencia *Oct4* y *Nanog* a día 14 de SCR mediante RT-qPCR. De manera consistente, se observó una reducción significativa para ambos marcadores después de la inhibición de *Dnmt2* en comparación al control transducido con shLuci (**Figura R1g**).

Posteriormente, nos propusimos investigar si la pérdida de función de DNMT2 tenía un impacto en la capacidad de proliferación celular. Los resultados no revelaron diferencias significativas entre los shRNAs dirigidos contra *Dnmt2* y el control shLuci, ni en los fibroblastos parentales (**Figura R2a**) ni en presencia de OSKM (**Figura R2b**). Esto sugiere que la menor eficiencia de reprogramación tras el silenciamiento de *Dnmt2* no está relacionada con defectos en la proliferación celular.

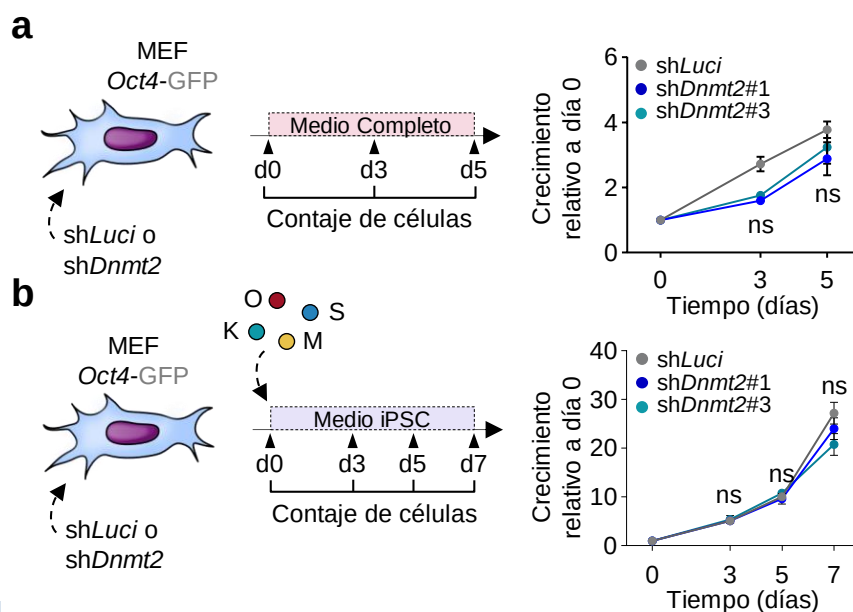


Figura R2. La falta de DNMT2 no afecta a la proliferación celular en MEFs. a) Esquema del análisis del efecto del silenciamiento de *Dnmt2* en el crecimiento de MEFs *Oct4*-GFP (izquierda) y cuantificación del crecimiento

relativo en comparación con el control (sh*Luci*) (derecha). El número de células se determinó en los días 0, 3 y 5 después de la transducción en medio completo para células somáticas. **b)** Esquema del análisis del crecimiento de MEFs *Oct4*-GFP tras la inhibición de *Dnmt2* en presencia de OSKM (izquierda) y cuantificación del crecimiento relativo en comparación con el control (sh*Luci*) (derecha). El número de células se determinó en los días indicados después de la transducción en medio de reprogramación (medio iPSC). En a y b, los datos se muestran como media $\pm$ desviación estándar de triplicados biológicos. Tras la realización de una prueba ANOVA de dos factores pareada con corrección de Dunnet para comparaciones múltiples, no se encontraron diferencias estadísticamente significativas, indicadas como ns: no significativo.

Finalmente, con el fin de comprobar que la inhibición de *Dnmt2* no solo retrasa la reprogramación celular, sino que la impide por completo, evaluamos el efecto de su inhibición tras 21 días en condiciones de reprogramación. Nuevamente, se transdujeron MEFs *Oct4*-GFP con shRNAs dirigidos contra *Dnmt2* o con el control sh*Luci*, junto con los factores de reprogramación OSKM (**Figura R3a**). Después de 21 días en medio de reprogramación, se cuantificó el número de colonias GFP⁺. De manera consistente con los resultados anteriores, encontramos una marcada reducción en el número de colonias GFP⁺ tras la inhibición de *Dnmt2* en comparación con el control transducido con sh*Luci* (**Figura R3b**). Así, nuestros resultados muestran que la inhibición de *Dnmt2* no retrasa el proceso de SCR, sino que lo impide por completo.

En conjunto, los datos obtenidos indican que la presencia de DNMT2 es necesaria para la generación de iPSCs.

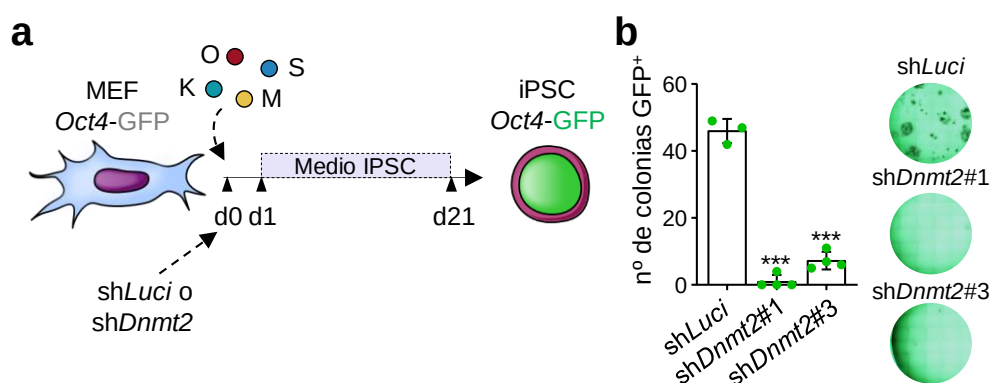


Figura R3. La pérdida de función de DNMT2 impide la SCR a largo plazo. **a)** Esquema de la estrategia de SCR seguida. Se llevó a cabo la transducción de MEFs *Oct4*-GFP con sh*Dnmt2* o sh*Luci*, como control, así como con los factores de reprogramación OSKM. Después de 21 días en presencia de medio de reprogramación, se determinó el número de colonias GFP⁺. **b)** Cuantificación del número de colonias GFP⁺ obtenidas en las condiciones indicadas (izquierda) e imágenes representativas de las colonias obtenidas (derecha). Se muestran los resultados como media $\pm$ desviación estándar y el número de replicados biológicos se indica por puntos. Para el análisis de diferencias estadísticamente significativas se aplicó una prueba ANOVA de un factor no pareada, con corrección de Dunnet para comparaciones múltiples. ***: $p < 0,001$.

6.1.2 La ausencia de DNMT2 impide la MET durante reprogramación

Dado que la reprogramación de fibroblastos a iPSCs implica una transición a través de múltiples identidades celulares permisivas (Plath y Lowry, 2011), nos propusimos investigar la existencia de una ventana temporal específica en la que la presencia de DNMT2 fuese requerida. Destacadamente, observamos que a los 21 días de la formación de iPSCs, la pérdida de función de DNMT2 no solo no logró inducir adecuadamente los niveles proteicos del marcador epitelial CDH1 (también conocido como E-CADHERINA), sino que tampoco fue capaz de suprimir los del marcador mesenquimal CDH2 (N-CADHERINA) (**Figura R4a**). Además, un análisis de expresión génica mediante RT-qPCR reveló que la inhibición de *Dnmt2* después de 14 días de SCR resultó en una reducción significativa de la activación

transcripcional de reguladores epiteliales (*Cdh1*, *Epcam*, *Cldn3* y *Ocln*) (**Figura R4b**) y en una ineficiente inhibición de reguladores mesenquimales (*Cdh2*, *Zeb1*, *Slug*, y *Vim*) (**Figura R4c**) en comparación con el control *shLuci*.

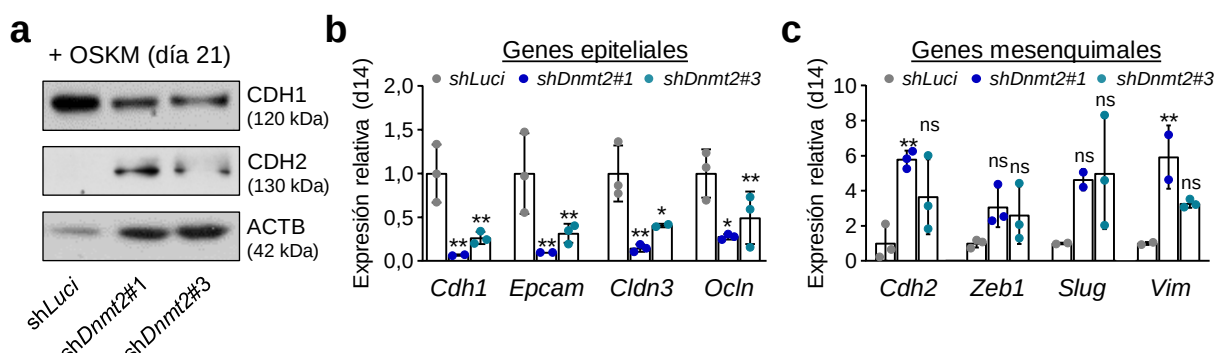


Figura R4. La inhibición de *Dnmt2* reduce la activación de programas transcripcionales epiteliales. a) Análisis de los niveles de las proteínas CDH1 y CDH2 mediante *western blot* (WB) después de 21 días de SCR utilizando OSKM en fibroblastos transducidos con *shDnmt2* o *shLuci* (control), con ACTB como control de carga. b) Análisis de los niveles de expresión de marcadores epiteliales mediante RT-qPCR después de la inhibición de *Dnmt2* a día 14 de SCR. c) Análisis de los niveles de expresión de marcadores mesenquimales, como en b. En b y c, los niveles de expresión se muestran relativos a los niveles de *U6* y al control (*shLuci*). Los datos se presentan como media $\pm$ desviación estándar de triplicados biológicos. Se aplicó una prueba ANOVA de un factor no pareada para calcular las diferencias estadísticamente significativas, con corrección de Dunnett para comparaciones múltiples. Estas se indican como: ns: no significativo, *: $p < 0,05$ y **: $p < 0,01$.

La activación de la expresión de marcadores epiteliales y la represión de marcadores mesenquimales son procesos que tienen lugar durante las primeras etapas de la SCR, mediante un proceso conocido como transición mesenquimal-epitelial (MET) (Li et al., 2010; Samavarchi-Tehrani et al., 2010). Teniendo esto en cuenta, los resultados obtenidos apuntan a un papel importante de DNMT2 en la supresión de los programas transcripcionales mesenquimales y en la activación de los epiteliales, lo que resulta necesario durante las etapas iniciales de la reprogramación de células somáticas mesenquimales en iPSCs. Para respaldar esta hipótesis, en primer lugar, realizamos un análisis *in silico* en el que se determinó cómo cambian los niveles de expresión de *Dnmt2* a lo largo del proceso de reprogramación de MEFs mediado por OSKM, utilizando datos de RNA-seq previamente publicados (Cieply et al., 2016). Aquí, se observó un ligero aumento de los niveles de *Dnmt2* al comienzo de la MET, seguido de su disminución después de esta etapa (**Figura R5a**). Estos patrones de expresión fueron validados mediante RT-qPCR en nuestro sistema de reprogramación (**Figura R5b**).

Ya que nuestros resultados apoyan un papel importante de DNMT2 en la activación de programas epiteliales y la supresión de programas mesenquimales durante la MET, decidimos investigar los cambios en la expresión génica que ocurren durante esta transición tras la inhibición de *Dnmt2*. Para abordar esto, se transdujeron MEFs *Oct4*-GFP con dos shRNAs independientes dirigidos contra *Dnmt2*, así como con el control (*shLuci*), junto con los factores de reprogramación OSKM. Después de 5 días en medio de reprogramación, realizamos un análisis del estado transcripcional celular mediante RNA-seq (**Figura R5c**). Previamente a la secuenciación masiva del transcriptoma, se validó el silenciamiento de *Dnmt2* en las células transducidas con cada uno de los shRNAs en comparación con el control (*shLuci*), a día 5 de SCR, mediante un análisis por RT-qPCR (**Figura R5d**).

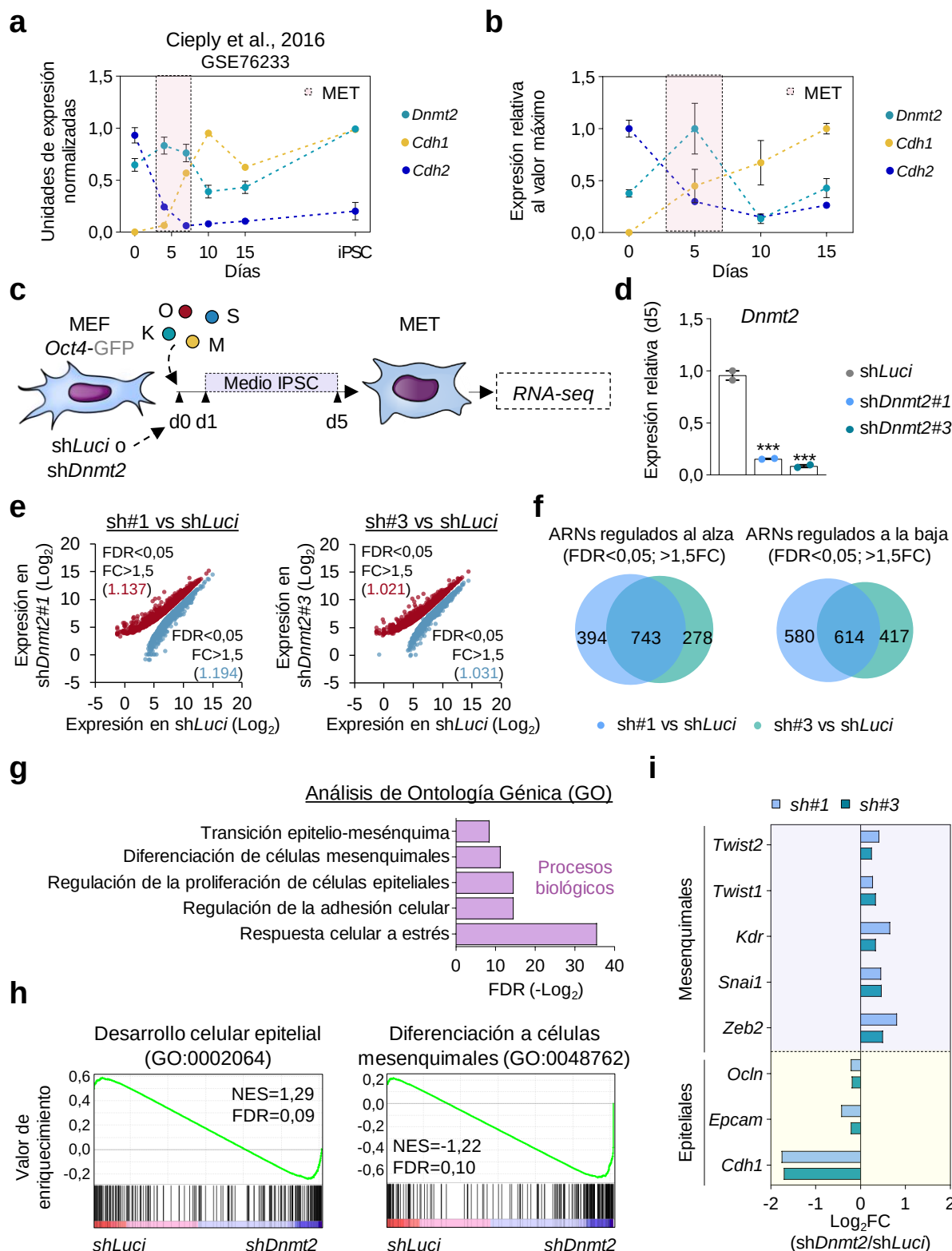


Figura R5. DNMT2 interviene en la activación del programa transcripcional epitelial al inicio de la SCR. a) Niveles de expresión de los genes *Cdh1*, *Cdh2* y *Dnmt2* durante la generación de iPSCs a partir de MEFs. Se analizaron los datos de RNA-seq del estudio de Cieply y colaboradores (GSE76233) (Cieply et al., 2016). b) Análisis de los niveles de expresión de *Cdh1*, *Cdh2* y *Dnmt2* mediante RT-qPCR en los días indicados de SCR. Los niveles de expresión se muestran relativos a los niveles de *Gapdh* y al valor máximo en cada gen analizado. Los datos se presentan como media $\pm$ desviación estándar de triplicados técnicos. c) Estrategia empleada para evaluar el transcriptoma mediante RNA-seq a día 5 de reprogramación mediada por OSKM a partir de MEFs *Oct4*-GFP

transducidos con shRNAs contra *Dnmt2* o *Luciferasa* (control). **d)** Análisis de los niveles de *Dnmt2* por RT-qPCR a día 5 de SCR en las condiciones indicadas. Los niveles se presentan relativos a los de *Hprt* y al control (sh*Luci*). Los datos se muestran como media $\pm$ desviación estándar y se indican mediante puntos el número de réplicas biológicas. Las diferencias estadísticamente significativas se evaluaron mediante una prueba ANOVA de un factor no pareada con corrección de Dunnet para comparaciones múltiples, ***: $p < 0,001$. **e)** Transcritos con regulación al alza (rojo) y a la baja (azul) en células con sh*Dnmt2*#1 (izquierda) y sh*Dnmt2*#3 (derecha) en comparación con el control (sh*Luci*), utilizando los criterios: $FDR < 0,05$ y $FC > 1,5$. **f)** Diagramas de Venn que muestran el número de transcritos regulados al alza (izquierda) y a la baja (derecha) tras la inhibición de *Dnmt2* con los distintos shRNAs. **g)** Análisis de ontología génica (GO) de los 1.357 (743+614) transcritos diferencialmente expresados después de la pérdida de función de DNMT2, mostrando los procesos biológicos relacionados. **h)** Análisis mediante enriquecimiento de conjuntos de genes (GSEA) de los datos de RNA-seq de células transducidas con sh*Dnmt2* (sh*Dnmt2*#1 y sh*Dnmt2*#3) o con sh*Luci* (control) a día 5 de SCR, empleando firmas génicas relacionadas con el desarrollo celular epitelial (GO:0002064) y la diferenciación hacia células mesenquimales (GO:0048762). Se muestran la puntuación de enriquecimiento normalizada (NES) y la tasa de descubrimientos falsos (FDR). **i)** Representación gráfica del cambio en la expresión ($\log_2FC$) de genes diferencialmente expresados ($FDR < 0,05$) relacionados con el fenotipo mesenquimal y epitelial en células transducidas con los diferentes shRNAs contra *Dnmt2* en comparación con el control (sh*Luci*) en el día 5 de la SCR.

A continuación, se examinaron las alteraciones transcripcionales causadas por la inhibición de *Dnmt2* con cada uno de los shRNAs empleados. Se seleccionaron aquellos transcritos cuya expresión variaba significativamente entre las condiciones analizadas (sh*Dnmt2* vs. sh*Luci*), con una tasa de descubrimientos falsos inferior al 5% ($FDR < 0,05$) y un cambio en la expresión mayor al 50% del valor en el control ($FC > 1,5$). Bajo estos criterios, se identificaron 1.194 y 1.031 transcritos con regulación negativa, así como 1.137 y 1.021 transcritos con regulación positiva, en células transducidas con sh*Dnmt2*#1 y sh*Dnmt2*#3, respectivamente (**Figura R5e**).

Con el fin de identificar el conjunto de genes cuya expresión se ve afectada de manera consistente por la inhibición de *Dnmt2*, se seleccionaron aquellos genes significativamente desregulados ($FDR < 0,05$ y $FC > 1,5$) en comparación con el control, detectados tanto en células transducidas con sh*Dnmt2*#1 como en células transducidas con sh*Dnmt2*#3 (**Figura R5f**). En este caso, la coincidencia entre los transcritos significativamente desregulados ($FDR < 0,05$, $FC > 1,5$) en células transducidas con sh*Dnmt2*#1 y sh*Dnmt2*#3 era mayor del 50% (**Figura R5f**). Finalmente, definimos una lista de 1.357 transcritos significativamente desregulados tras la pérdida de función de *Dnmt2*, de los cuales 743 se encontraban regulados al alza y 614 regulados a la baja (**Figura R5f**).

Posteriormente, se llevó a cabo un análisis de ontología génica con el objetivo de esclarecer la relevancia biológica de la pérdida de función de DNMT2 en la MET durante la reprogramación celular. Para este propósito, empleamos la herramienta en línea DAVID (Huang da et al., 2009; Sherman et al., 2022). Este análisis reveló que los genes diferencialmente expresados tras la inhibición de *Dnmt2* en etapas iniciales del proceso de SCR participaban en procesos clave de determinación del destino celular, como la transición epitelial-mesenquimal (EMT), la adhesión celular, la proliferación de células epiteliales o la diferenciación hacia células mesenquimales (**Figura R5g**).

En sintonía con estos hallazgos, un análisis de enriquecimiento de conjuntos de genes (GSEA) reveló que la expresión de genes implicados en el desarrollo celular epitelial (GO:0002064) se encuentra enriquecida en las muestras control, respecto a las que presentan bajos niveles de expresión de *Dnmt2* (**Figura R5h**, panel de la izquierda). Por otro lado, la expresión de genes relacionados con la diferenciación hacia células mesenquimales (GO:0048762) se encuentra enriquecida en las células donde *Dnmt2* es silenciado en comparación con las células control (**Figura R5h**, panel de la derecha). Estos resultados respaldan que la pérdida de función de *Dnmt2* ocasiona defectos, a nivel transcripcional, en la

adquisición de un destino celular epitelial durante la MET. De acuerdo con esta hipótesis, se determinó una regulación al alza de varios marcadores mesenquimales, como *Twist1*, *Twist2*, *Kdr*, *Snai1* y *Zeb2*, entre otros (**Figura R5i**). Asimismo, se constató una regulación a la baja de reguladores epiteliales como *Cdh1*, *Ocln* y *Epcam*, en células transducidas, tanto con *shDnmt2#1*, como con *shDnmt2#3*, en comparación con las células transducidas con *shLuci* (**Figura R5i**).

Finalmente, después de 7 días de reprogramación mediada por OSKM (**Figura R6a**), confirmamos que la ausencia de DNMT2 (**Figuras R6b y R6c**) resultaba en una marcada reducción del principal regulador epitelial CDH1, tanto a nivel transcripcional, como de proteína (**Figuras R6d y R6e**). Asimismo, los niveles del marcador mesenquimal CDH2 se mantuvieron elevados en las células transducidas con *shDnmt2* en comparación con el control *shLuci* (**Figuras R6f y R6g**).

Estos resultados, sumados a los previamente obtenidos, señalan que la inhibición de la expresión de *Dnmt2* conlleva un defecto en la activación y represión de programas epiteliales y mesenquimales, respectivamente, impidiendo un correcto desarrollo del proceso de MET en las primeras etapas de la SCR.

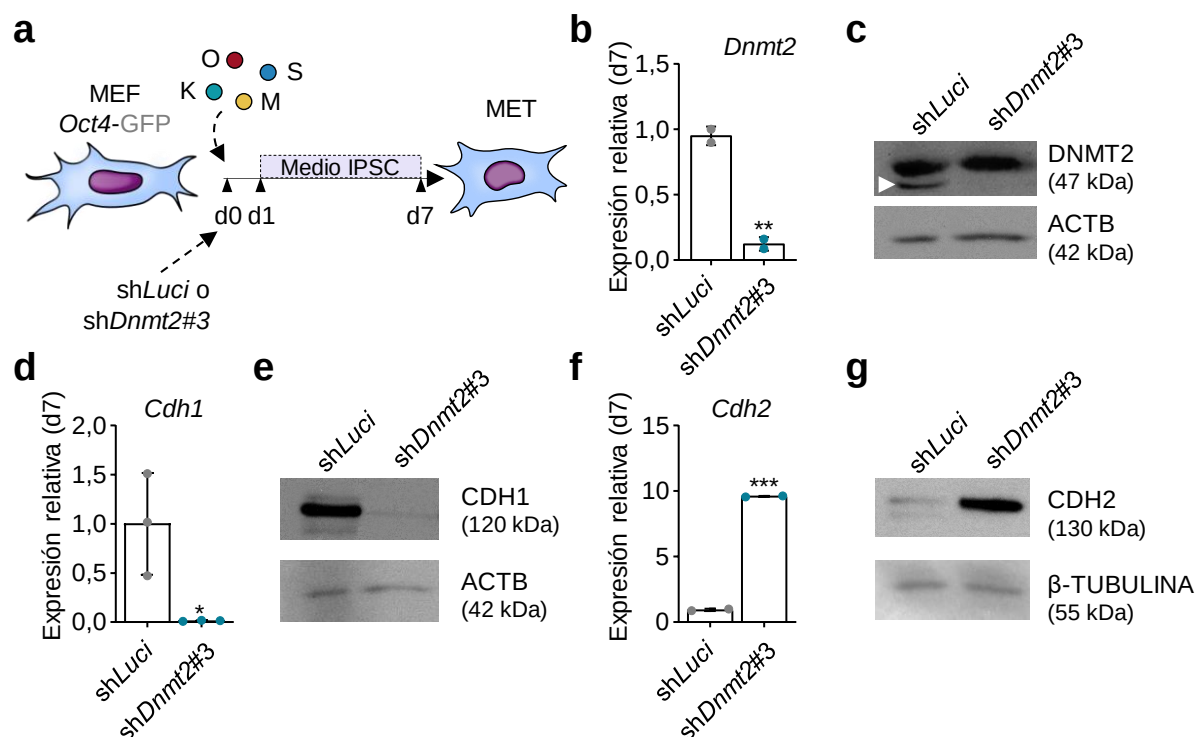


Figura R6. El silenciamiento de *Dnmt2* afecta a la activación de *Cdh1* y a la represión de *Cdh2* en las primeras etapas de la SCR. **a**) Estrategia de reprogramación celular. Los MEFs *Oct4-GFP* se transdujeron con *shDnmt2#3* o *shLuci* junto con los factores OSKM. Tras 7 días de reprogramación, se evaluaron los niveles de ARN mediante RT-qPCR (paneles **b**, **d** y **f**) y de proteína mediante WB (paneles **c**, **e** y **g**) de los genes *Dnmt2*, *Cdh1* y *Cdh2*, respectivamente. En **b**, **d** y **f**, se presentan los niveles de expresión relativos a los de *Hprt* y al control (*shLuci*). Los datos se muestran como media $\pm$ desviación estándar y se indican mediante puntos el número de réplicas biológicas. Se utilizaron pruebas t de Student no pareadas de una (**b**) y dos colas (**d**, **f**), para evaluar las diferencias estadísticamente significativas, indicadas como *: $p < 0,05$, **: $p < 0,01$ y ***: $p < 0,001$. En **c**, **e** y **g**, ACTB (**c**, **e**) o β -TUBULINA (**g**) se utilizaron como controles de carga. En **c**, la cabeza de flecha señala la banda específica correspondiente a DNMT2.

6.1.3 La función catalítica de DNMT2 es necesaria durante la MET

Una vez confirmado el papel crítico de DNMT2 durante la MET en las primeras etapas de la SCR, nos propusimos determinar si esta función es dependiente o no de su actividad catalítica.

En primer lugar, analizamos los niveles de metilación en la citosina 38 (m^5C_{38}) del ARNt que porta el ácido aspártico (ARNt-Asp) durante el proceso de SCR, ya que constituye una de las principales dianas descritas hasta el momento de DNMT2 (Goll et al., 2006; Schaefer et al., 2010; Tuorto et al., 2015; Zhang et al., 2018). Para ello, se extrajo ARN total de fibroblastos transducidos con *shLuci* o *shDnmt2* en distintas etapas clave del proceso de generación de iPSCs, que incluyeron: la fase inicial a los 5 días, la fase de maduración a los 10 días y la fase de estabilización a los 15 días. Tras el tratamiento del ARN con bisulfito para identificar las posiciones modificadas con m^5C , realizamos la retrotranscripción y amplificación específica del ARNt-Asp. Luego, clonamos esta molécula en un vector para su secuenciación por Sanger. Los resultados indicaron que en las células control (*shLuci*), los niveles de m^5C en la posición C38 (m^5C_{38}) del ARNt-Asp se mantienen elevados durante la SCR, disminuyendo significativamente tras la inhibición de *Dnmt2* (**Figura R7a**). Esto sugiere que la función catalítica de DNMT2 podría resultar esencial para la reprogramación celular, de modo que su inhibición comprometería la obtención de células pluripotentes.

Para confirmar esta hipótesis, validamos en primer lugar que la expresión forzada de DNMT2 permitía rescatar la capacidad de reprogramación celular tras su inhibición, y que por lo tanto, se trataba de un efecto directo de la pérdida de esta proteína. Se transdujeron fibroblastos adultos, derivados de orejas de ratones de 5 meses de edad, con *shLuci* o *shDnmt2#3*, específicamente dirigido contra la región 3'UTR no codificante de este gen. En las células silenciadas para *Dnmt2*, se llevó a cabo la expresión forzada de la proteína humana DNMT2 ($hDNMT2^{WT}$), previamente clonada en un sistema retroviral de expresión constitutiva (pMX) (**Figura R7b**). Paralelamente, también se llevó a cabo la transducción con los factores de reprogramación OSKM (**Figura R7b**) y tras 15 días en condiciones de reprogramación, se procedió a el conteo de colonias con morfología pluripotente para cada condición (**Figura R7c**). Tal y como habíamos observado previamente, la inhibición de *Dnmt2* resultó en una reducción drástica del número de iPSCs obtenidas, mientras que la sobreexpresión de la proteína DNMT2 humana permitió alcanzar una tasa de reprogramación similar a las células transducidas con *shLuci* (**Figura R7c**). De este modo, estos resultados confirman que la disrupción de la capacidad de reprogramación celular observada tras la inhibición de *Dnmt2* se debe a un efecto directo de la pérdida de esta proteína.

Luego, nos propusimos evaluar si DNMT2 desempeña su función en la generación de iPSCs a través de su actividad catalítica sobre el ARN. Para ello, los fibroblastos adultos fueron transducidos con *shDnmt2#3* y a su vez con el constructo que permitía la expresión constitutiva de $hDNMT2^{WT}$ o de una versión inactiva para la deposición de la marca m^5C en ARN ($hDNMT2^{C79A}$) (Jurkowski et al., 2008). Es relevante mencionar que los vectores que contenían las secuencias codificantes para las proteínas humanas DNMT2 WT y mutantes (entre las que se encuentra la proteína $hDNMT2^{C79A}$), fueron proporcionados por el laboratorio del Dr. Albert Jeltsch (Universidad de Stuttgart), para su posterior clonaje en vectores retrovirales de expresión constitutiva. De la misma manera que en ensayos previos, se llevó a cabo la transducción con *shLuci* como control. La validación de la actividad catalítica de cada constructo se llevó a cabo en ESCs (consultar la sección 6.2.4). Paralelamente, estas células se transdujeron con los factores de reprogramación OSKM y se mantuvieron en condiciones de

reprogramación durante 15 días (**Figura R7d**). Finalmente, con el objetivo de determinar si la actividad catalítica de DNMT2 es necesaria durante la generación de iPSCs, y en particular durante la MET, se realizó el análisis de la expresión de marcadores epiteliales (*Cdh1* y *Ocln*) y de pluripotencia (*Oct4* endógeno). Los resultados obtenidos indicaron que la reducción en la expresión de los genes *Cdh1*, *Ocln* y *Oct4* después de la pérdida de función de DNMT2 se pudo restaurar a niveles comparables al control al forzar la expresión de la forma catalíticamente activa (WT), pero no de la versión mutante catalítica (C79A) de DNMT2 (**Figura R7e**).

En conjunto, estos resultados muestran que la actividad catalítica de DNMT2 sobre el ARN resulta esencial durante MET, la fase más temprana de la reprogramación de fibroblastos a iPSCs.

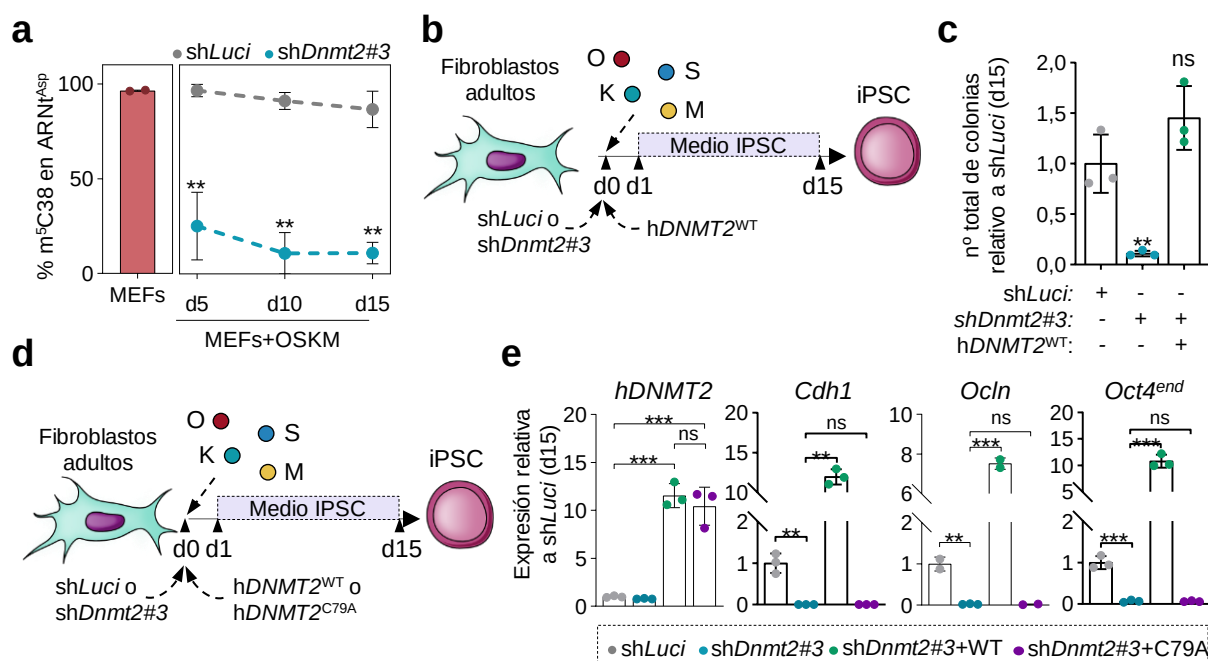


Figura R7. La pérdida de actividad catalítica de DNMT2 en ARN compromete la generación de iPSCs. a) Análisis de los niveles de metilación en la posición C38 de ARN^{Asp} en MEFs y en los días indicados de SCR tras la inhibición de *Dnmt2* o de *Luciferasa* (control), utilizando secuenciación Sanger de muestras de ARN tratadas con bisulfito. Los datos se expresan como media ± desviación estándar de dos y tres réplicas biológicas en MEFs y durante SCR, respectivamente. Se realizó un análisis estadístico mediante pruebas t de Student para cada punto temporal, con corrección FDR para comparaciones múltiples. **: p < 0,01. b) Estrategia de reprogramación empleada. Se transdujeron fibroblastos adultos con shDnmt2#3 o shLuci (control). Además, después de la inhibición de *Dnmt2*, se sobreexpresó la proteína humana DNMT2 (hDNMT2^{WT}). Después de 15 días en condiciones de reprogramación, se determinó el número de colonias pluripotentes. c) Representación del número relativo de iPSCs en cada condición indicada en comparación con el control (shLuci) a los 15 días de SCR. d) Estrategia de reprogramación mediada por OSKM empleada a partir de fibroblastos adultos con shDnmt2#3 o shLuci (control). En este caso, después de la inhibición de *Dnmt2*, se sobreexpresó hDNMT2^{WT} o su variante mutante catalítica hDNMT2^{C79A}. e) Análisis de los niveles de expresión de los marcadores epiteliales *Cdh1* y *Ocln*, así como de la expresión endógena del marcador de pluripotencia *Oct4* (*Oct4*^{end}), en células transducidas con shLuci, shDnmt2#3, shDnmt2#3 con hDNMT2^{WT} o shDnmt2#3 con hDNMT2^{C79A}, a los 15 días de SCR. Se presentan los niveles de expresión relativos a los de *Hprt* y al control (shLuci). En c y e, los datos se presentan como media ± desviación estándar y se indica mediante puntos el número de réplicas biológicas. Se evaluaron las diferencias estadísticas mediante una prueba ANOVA de un factor no pareada, con corrección de Sidak para comparaciones múltiples, indicadas como ns: no significativo, **: p < 0,01 y ***: p < 0,001.

Dado que los datos obtenidos hasta el momento indican que la actividad catalítica de DNMT2 es crucial en el proceso de MET, decidimos investigar si la sobreexpresión de esta actividad podría facilitar el proceso. Para esto, llevamos a cabo la sobreexpresión de hDNMT2^{WT} o de la versión mutante catalítica hDNMT2^{C79A} en MEFs *Oct4*-GFP, de manera

inducible. Ambas proteínas habían sido previamente clonadas en un sistema lentiviral de expresión inducible (FUW-Tet-On), que es activado en presencia del transactivador M2rtTA, tras la adición de doxiciclina (DOX) (Urlinger et al., 2000). Paralelamente, transdujimos los MEFs con los factores OSKM y después de 14 días en medio de reprogramación suplementado con DOX (2 $\mu\text{g}/\mu\text{L}$), evaluamos el número de colonias GFP+ (**Figura R8a**). En primer lugar, evaluamos la eficacia de la sobreexpresión del gen *DNMT2* humano mediante RT-qPCR, detectándose un marcado aumento en la expresión de ambas formas, *hDNMT2*^{WT} y *hDNMT2*^{C79A}, en comparación con el control negativo (M2rtTA+DOX) (**Figura R8b**). Los resultados demostraron un incremento en el número total de colonias iPSC en comparación con el control, tanto atendiendo a su morfología, como a los niveles de GFP. Este aumento fue observado tras la sobreexpresión de la forma catalíticamente activa, pero no de la mutante (**Figuras R8c y R8d**), apoyando nuestros resultados anteriores que sugieren que DNMT2 ejerce su función durante la SCR a través de su actividad catalítica.

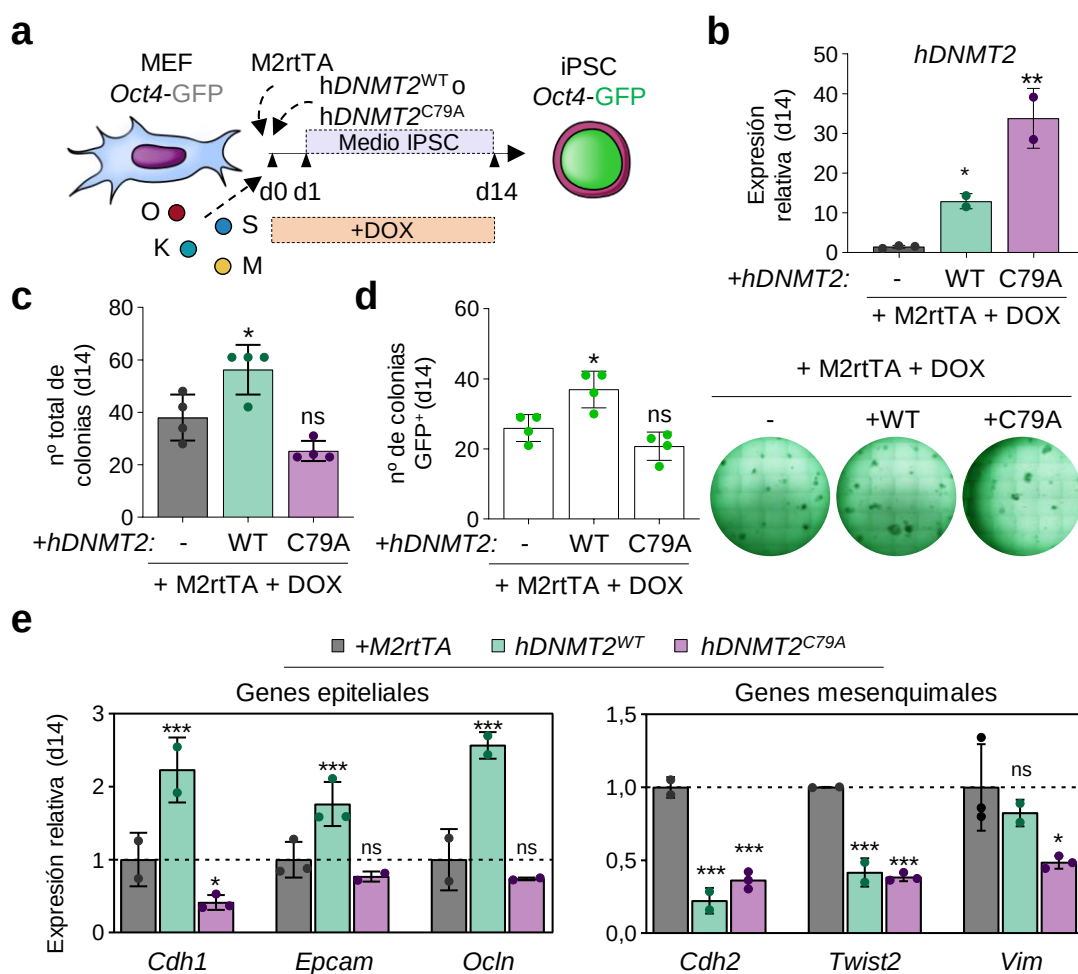


Figura R8. La ganancia de función de la actividad catalítica de DNMT2 favorece la generación de iPSCs. a) Estrategia de reprogramación empleada. Los MEFs Oct4-GFP se transdujeron con constructos inducibles FUW-Tet-On que permiten la expresión de *hDNMT2*^{WT} o *hDNMT2*^{C79A} en presencia del transactivador M2rtTA y doxiciclina (DOX). Tras la transducción con los factores de reprogramación OSKM, las células se mantuvieron en medio de reprogramación y en presencia de DOX (2 $\mu\text{g}/\mu\text{L}$) durante 14 días. b) Análisis de la expresión de *hDNMT2* mediante RT-qPCR a los 14 días de la reprogramación. c) Cuantificación mediante microscopía de campo claro del número de colonias con morfología iPSC obtenidas a partir de las condiciones indicadas a día 14 de SCR. d) Cuantificación mediante microscopía de fluorescencia del número de colonias GFP+ en las condiciones indicadas (izquierda) e imágenes representativas de las colonias obtenidas (derecha). e) Niveles de expresión, evaluados por RT-qPCR, de marcadores epiteliales (izquierda) y marcadores mesenquimales (derecha), después de la sobreexpresión de la forma WT o mutante catalítica (C79A) de *hDNMT2* a día 14 de SCR. En b y e, se muestran

los niveles de expresión relativos a los de *Actb* y al control (M2rtTA + DOX). En b, c, d y e, se muestran los resultados como media $\pm$ desviación estándar y el número de replicados biológicos se indica por puntos. Las diferencias estadísticas se calcularon utilizando una prueba ANOVA de un factor no pareada con corrección de Sidak para comparaciones múltiples y se indican como ns: no significativo, *: $p < 0,05$, **: $p < 0,01$ y ***: $p < 0,001$.

A continuación, para confirmar que la metilación del ARN por parte de DNMT2 promueve el proceso de MET, analizamos la expresión de marcadores epiteliales (*Cdh1*, *Epcam* y *Ocln*) y mesenquimales (*Cdh2*, *Twist2* y *Vim*), después de 14 días de reprogramación celular (**Figura R8e**). Los resultados obtenidos revelaron un aumento significativo en la expresión de los marcadores epiteliales bajo la sobreexpresión de la forma catalíticamente activa, pero no de la mutante (**Figura R8e**). Sin embargo, también se observó una disminución en los niveles de los marcadores mesenquimales después de la sobreexpresión de ambas formas, tanto la activa como la mutante catalítica (**Figura R8e**).

Estos resultados, junto con los obtenidos previamente, evidencian que la actividad catalítica de DNMT2 es esencial para la activación de programas transcripcionales epiteliales durante el proceso de MET, lo cual es necesario para la generación de iPSCs a partir de células somáticas mesenquimales. Sin embargo, esta actividad no parece estar involucrada en la supresión de programas transcripcionales mesenquimales. Esto podría deberse a que funciones catalítico dependientes e independientes de la proteína controlen la expresión de grupos distintos de transcritos.

6.1.4 La función de DNMT2 es prescindible en etapas avanzadas de la reprogramación celular somática

Dado que nuestros hallazgos indican que la función de DNMT2 es crucial durante la MET, nos propusimos investigar si dicha actividad sigue siendo necesaria en etapas posteriores del proceso de SCR, una vez que el destino celular epitelial ya ha sido establecido.

Para abordar esta cuestión, llevamos a cabo un experimento de reprogramación celular utilizando células intermediarias avanzadas de reprogramación, comúnmente conocidas como "pre-iPSCs" (Plath y Lowry, 2011). En este caso, generamos una línea clonal de pre-iPSCs a partir de MEFs con un reportero *Nanog*-GFP. Estos MEFs fueron previamente transducidos con los factores OSKM para iniciar la reprogramación. Tras 12 días de exposición a condiciones de reprogramación en medio estándar con suero y LIF (medio SL), las pre-iPSCs se seleccionaron utilizando citometría de flujo activada por fluorescencia (FACS) para obtener una población de células GFP-. Luego, estas células GFP-, que no eran propensas a la reprogramación en presencia de SL, se transfectaron con constructos PiggyBac que codificaban *Nanog* o un vector vacío (EV). Ambas líneas celulares se mantuvieron en condiciones de cultivo estándar y se transdujeron con shRNAs dirigidos contra *Dnmt2* o *Luciferasa*, como control. Después de la selección con antibióticos para garantizar la expresión estable de los shRNAs, se inició la reprogramación al cambiar el medio de SL a uno sin suero que contenía LIF y los inhibidores de las quinasas MEK y GSK3 β (medio 2iL) (**Figura R9a**). Los resultados revelaron, en primer lugar, que las pre-iPSCs que sobreexpresaban *Nanog* generaron diez veces más colonias GFP+ en comparación con aquellas que recibieron el EV (**Figuras R9b y R9c**), en consonancia con estudios previos (Costa et al., 2013; Guallar et al., 2020; Theunissen et al., 2011). Además, demostraron que la pérdida de la función de DNMT2 no afectó la generación de colonias GFP+ y no alteró significativamente la eficiencia de la reprogramación, ya sea en presencia o ausencia de NANOG (**Figuras R9b y R9c**).

Estos resultados indican que la función de DNMT2 no es requerida en etapas posteriores de la reprogramación de las iPSCs, una vez que la identidad celular epitelial ya ha sido establecida. Además, evidencian su relevancia en las etapas tempranas de la reprogramación, es decir, en la transición mesenquimal-epitelial (MET), para la formación exitosa de iPSCs.

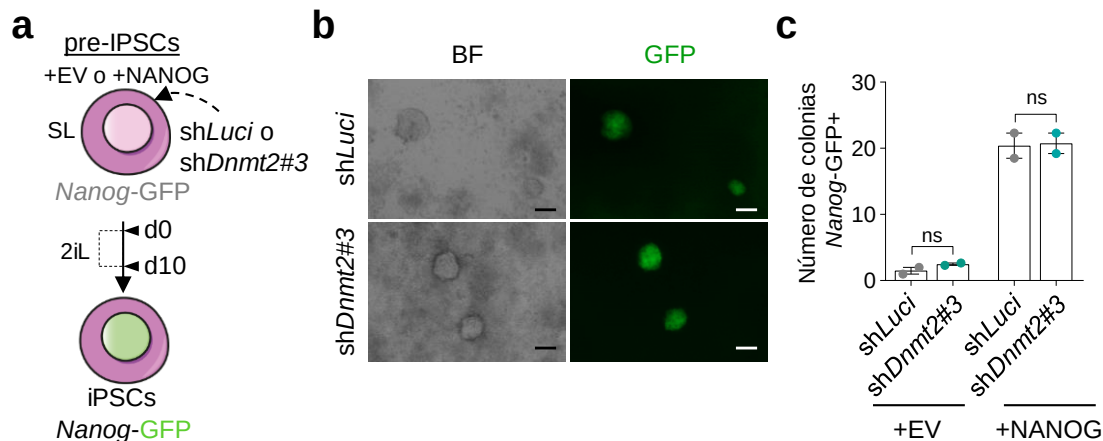


Figura R9. La ausencia de DNMT2 no afecta a la reprogramación de intermediarios epiteliales a iPSCs. a) Estrategia de reprogramación de intermediarios epiteliales avanzados (pre-iPSCs) seguida. Las pre-iPSCs *Nanog*-GFP con sobreexpresión estable de *Nanog* o de un vector vacío (EV) en medio SL fueron transducidas con shRNAs dirigidos contra *Dnmt2* o *Luciferasa* (shLuci), como control. Tras 10 días en cultivo en medio de reprogramación 2iL, se cuantificó el número de colonias GFP+ mediante microscopía de fluorescencia. b) Imágenes representativas en campo claro (BF, izquierda) y de fluorescencia (GFP, derecha) de las colonias GFP+ obtenidas en las distintas condiciones de estudio. Escala: 100 μ m. c) Cuantificación de las colonias GFP+ después de 10 días de reprogramación en medio 2iL en las condiciones indicadas. Los datos se muestran como media $\pm$ desviación estándar de dos replicados biológicos. Las diferencias estadísticamente significativas se calcularon mediante pruebas t de Student de dos colas no pareadas con corrección FDR para comparaciones múltiples y se indican como ns: no significativo.

6.1.5 DNMT2 interviene en la reprogramación celular directa a precursores hematopoyéticos

Teniendo en cuenta el papel fundamental de DNMT2 durante la reprogramación hacia células pluripotentes, nos planteamos la posibilidad de que su función fuese también necesaria en otros procesos que implicasen cambios de identidad celular. Para evaluar esta posibilidad, utilizamos un sistema de reprogramación directa (también denominado transdiferenciación), que permite la conversión de células somáticas diferenciadas en precursores hematopoyéticos. Esta aproximación experimental fue llevada a cabo durante mi estancia predoctoral de tres meses en el laboratorio del Dr. Carlos Filipe Pereira, en la Universidad de Lund (Suecia). El Dr. Pereira fue un pionero en el desarrollo de este sistema, que se basa en la expresión forzada de tres factores de transcripción (GATA2, GF11b y FOS) y que permite la conversión de fibroblastos murinos y humanos en precursores hematopoyéticos inducidos (iHSCs) (Pereira et al., 2016; Pereira et al., 2013).

Para explorar la función de DNMT2 en la generación de precursores hematopoyéticos, inicialmente, aplicamos una estrategia de pérdida de función en MEFs que portaban el marcador H2B-GFP bajo el control del promotor del gen humano *CD34*, cuya expresión es característica de las células precursoras hematopoyéticas y disminuye a medida que éstas se diferencian (Radomska et al., 2002). En las iHSCs, la proteína GFP se encuentra fusionada a la histona H2B, lo que permite la visualización de los núcleos celulares que la expresan (Pereira et al., 2013; Schaniel y Moore, 2009). Tras la transducción de los MEFs *CD34*/H2BGFP con shRNAs contra *Luciferasa* (control) o *Dnmt2*, que fueron posteriormente seleccionados durante dos días

para garantizar la expresión estable de los constructos, procedimos a la sobreexpresión de los factores de reprogramación GATA2, GFI1b y FOS (denominados colectivamente como GGF), lo cual se llevó a cabo durante dos días consecutivos y utilizando vectores lentivirales FUW de expresión constitutiva, facilitados en el laboratorio del Dr. Pereira. Seguidamente, las células fueron mantenidas durante 15 días en medio de cultivo hematopoyético (mielocultivo) y después, se procedió a evaluar la eficiencia de reprogramación mediante microscopía de fluorescencia y citometría de flujo (**Figura R10a**). Los resultados obtenidos revelaron una reducción importante de los niveles de H2B-GFP tras la inhibición de *Dnmt2*, tanto mediante visualización por microscopía de fluorescencia (**Figura R10b**), como por citometría de flujo, donde se detectó una reducción de aproximadamente el 50% en la eficiencia de reprogramación hacia iHSCs (**Figura R10c**).

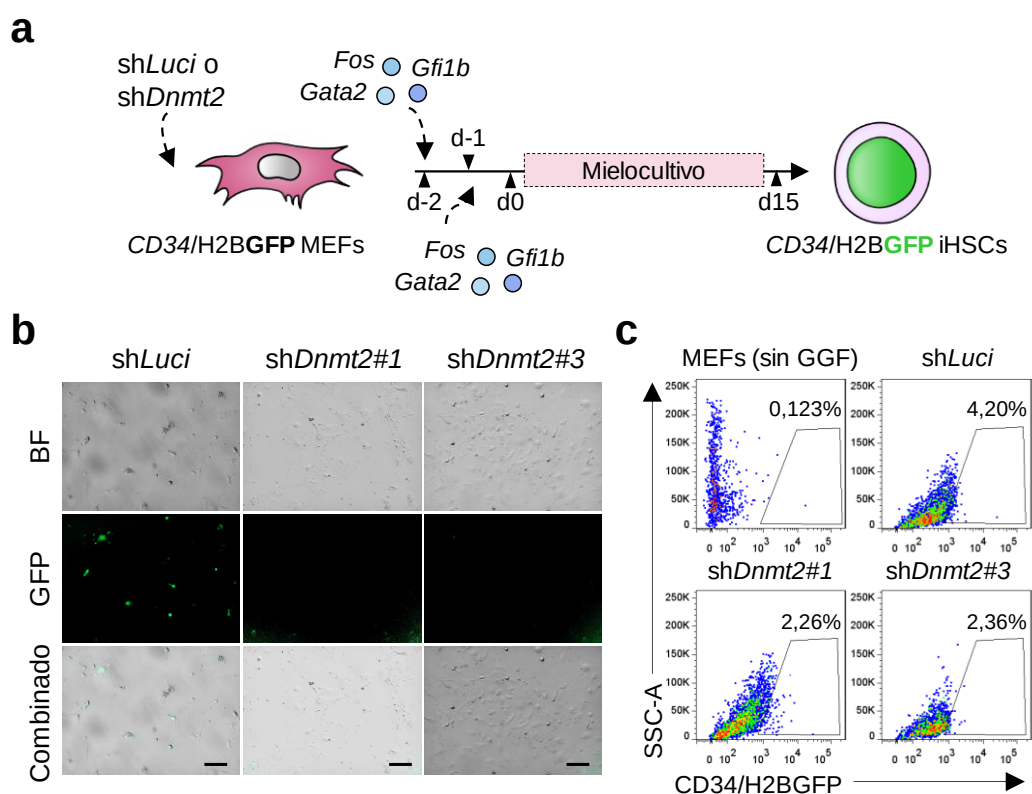


Figura R10. La inhibición de DNMT2 dificulta la obtención de precursores hematopoyéticos a partir de MEFs. **a)** Esquema de la estrategia de reprogramación utilizada. Los MEFs CD34/H2BGFP se transdujeron con *shDnmt2* o *shLuci* y después de asegurar la expresión estable de los constructos mediante selección con antibiótico, se indujo la sobreexpresión de los factores de transcripción GATA2, GFI1b y FOS (GGF) durante dos días consecutivos. Luego, las células se mantuvieron en medio de cultivo hematopoyético (mielocultivo) durante 15 días y se evaluó la presencia de células GFP+ mediante microscopía y citometría de flujo. **b)** Imágenes representativas en campo claro (BF, panel superior), de fluorescencia (GFP, panel del medio) y la superposición de ambas (panel inferior) a los 15 días de reprogramación hacia iHSCs en células transducidas con *shDnmt2* o *shLuci*. Escala: 320 μ m. **c)** Representación del porcentaje de células GFP+ obtenidas por citometría de flujo tras la transducción con *shDnmt2* o *shLuci* a los 15 días de reprogramación hacia iHSCs. Se muestran CD34/H2BGFP MEFs no reprogramados como control negativo.

Dada la función mostrada de DNMT2 durante la reprogramación de MEFs hacia iPSCs, así como en la transdiferenciación de MEFs hacia iHSCs, nos propusimos investigar si esta función se encuentra conservada entre especies. Para abordar esta cuestión, evaluamos si la supresión de *DNMT2* tenía un efecto negativo en la eficiencia de la reprogramación directa de fibroblastos dérmicos humanos (HDFs) hacia precursores hematopoyéticos siguiendo el

protocolo optimizado por el grupo del Dr. Pereira para este propósito (Silverio-Alves et al., 2019) (**Figura R11a**).

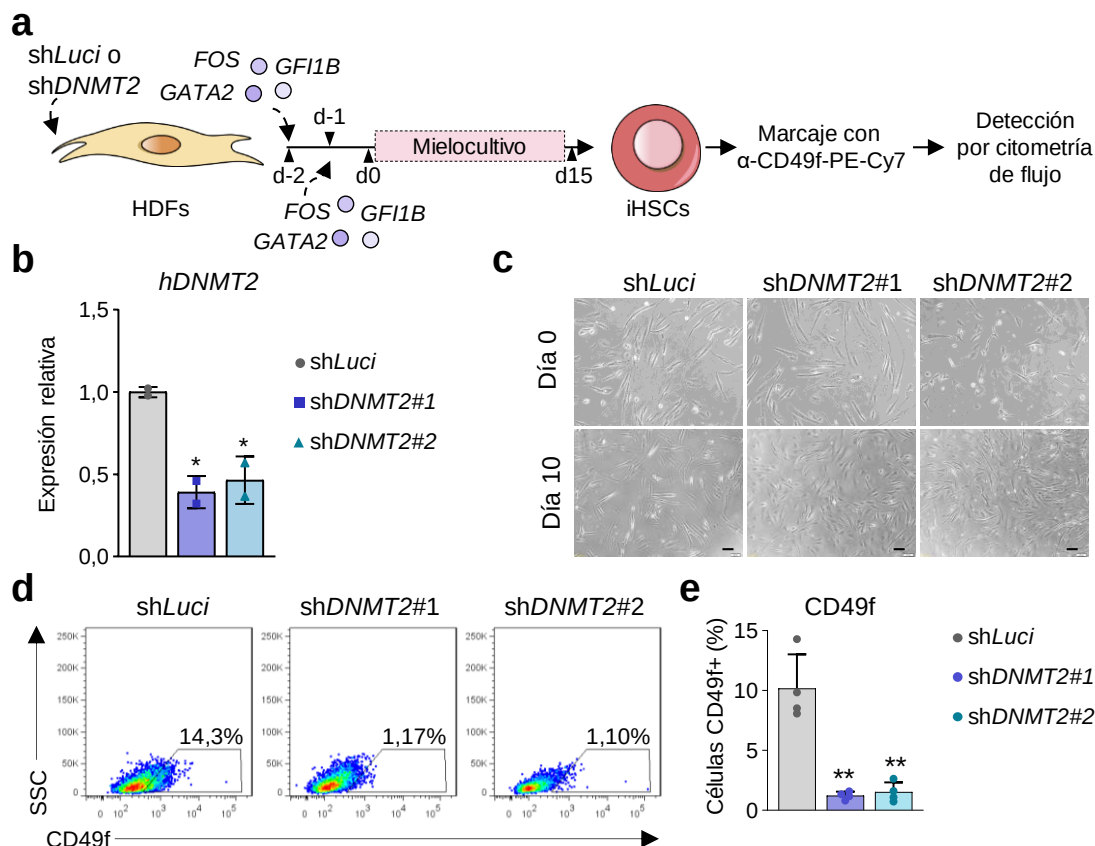


Figura R11. La función de DNMT2 es necesaria para la generación eficiente de iHSCs humanas. a) Esquema de la estrategia de reprogramación utilizada. Los HDFs se transdujeron con shRNAs independientes dirigidos contra el gen humano *DNMT2* o con shLuci (control). Luego de asegurar la expresión estable de los constructos mediante selección con 2 µg/mL del antibiótico puromicina, se indujo la expresión de los factores de reprogramación humanos GGF en dos días consecutivos. A continuación, las células se mantuvieron en medio de cultivo hematopoyético durante 15 días y se procedió al marcaje de las células con un anticuerpo contra CD49f conjugado con PE-Cy7. La señal se detectó mediante citometría de flujo. b) Análisis de los niveles de expresión del gen humano *DNMT2* en HDFs transducidos con shLuci o con shRNAs contra *DNMT2*. Los niveles de expresión se relativizaron a los de *GAPDH* y al control (shLuci). c) Imágenes representativas en microscopía de campo claro, a los días 0 (panel superior) y 10 (panel inferior), del proceso de reprogramación hacia iHSCs en células transducidas con shRNAs dirigidos contra *DNMT2* o con shLuci (control). Escala: 200 µm. d) Porcentaje de células CD49f+ detectadas mediante citometría de flujo a los 15 días de reprogramación hacia iHSCs en las condiciones indicadas. e) Cuantificación del porcentaje de células CD49f+ obtenidas por citometría de flujo a partir de HDFs transducidos con shLuci o con shRNAs contra *DNMT2*, después de 15 días de reprogramación hacia iHSCs. En b y e, los datos se muestran como media ± desviación estándar y el número de replicados biológicos se indica por puntos. Para el cálculo de diferencias estadísticamente significativas, se aplicó una prueba ANOVA de un factor no pareada con corrección de Dunnett para comparaciones múltiples. Estas se indican como *: p<0,05 y **: p<0,01.

En primer lugar, diseñamos y clonamos dos shRNAs independientes dirigidos contra la versión humana del gen *DNMT2*. Luego, verificamos la eficacia del silenciamiento por RT-qPCR en HDFs transducidos con cada uno de estos shRNAs en comparación con el control shLuci (**Figura R11b**). Después, aseguramos la expresión estable de los shRNAs mediante la selección con antibióticos. En los dos días siguientes, se indujo la expresión de los factores de transcripción humanos GATA2, GFI1B y FOS (GGF). Finalmente, las células se mantuvieron durante 15 días en un medio de cultivo hematopoyético diseñado específicamente para la reprogramación de células humanas hacia iHSCs (**Figura R11a**).

Para evaluar la eficiencia de reprogramación, se procedió al marcaje de las células con un anticuerpo fluorescente dirigido al marcador de membrana CD49f, característico del endotelio hematopoyético y luego se examinó su presencia mediante citometría de flujo (**Figura R11a**). Es importante destacar que, durante el proceso de reprogramación no se detectaron defectos evidentes en el crecimiento de los fibroblastos, como se pudo apreciar mediante microscopía visible a lo largo de los experimentos realizados (**Figura R11c**). De manera consistente con los resultados obtenidos en el sistema de transdiferenciación murino, observamos que la reducción de los niveles de *DNMT2* fue suficiente para generar una notable disminución en los niveles del marcador de endotelio hematopoyético CD49f durante el proceso de reprogramación directa hacia iHSCs a partir de fibroblastos humanos (**Figuras R11d y R11e**).

Después de observar que la inhibición de *DNMT2* conlleva una reducción en la eficiencia de reprogramación de fibroblastos a precursores hematopoyéticos, nos planteamos si el aumento de su función podría mejorar la obtención de iHSCs. Con este propósito, llevamos a cabo la sobreexpresión del gen *Dnmt2* murino en *CD34/H2BGFP* MEFs utilizando un vector lentiviral de expresión inducible (FUW-Tet-On-*Dnmt2*), en presencia del transactivador M2rtTA (Urlinger et al., 2000). Después, indujimos la expresión de los factores de transcripción GGF en dos días consecutivos y mantuvimos las células durante 15 días en un medio de cultivo hematopoyético para células murinas (**Figura R12a**).

Posteriormente, evaluamos la presencia de colonias de iHSCs mediante microscopía de fluorescencia y citometría de flujo. En este caso, observamos un aumento en el número de células reprogramadas después de la sobreexpresión de *Dnmt2* en comparación con el control con solo la presencia del transactivador M2rtTA y DOX (**Figura R12b**). Además, el análisis por citometría de flujo reveló un notable incremento en el número de células GFP+ después de la sobreexpresión de *Dnmt2* (5,93%) en comparación con el control (1,77%) (**Figuras R12c y R12d**).

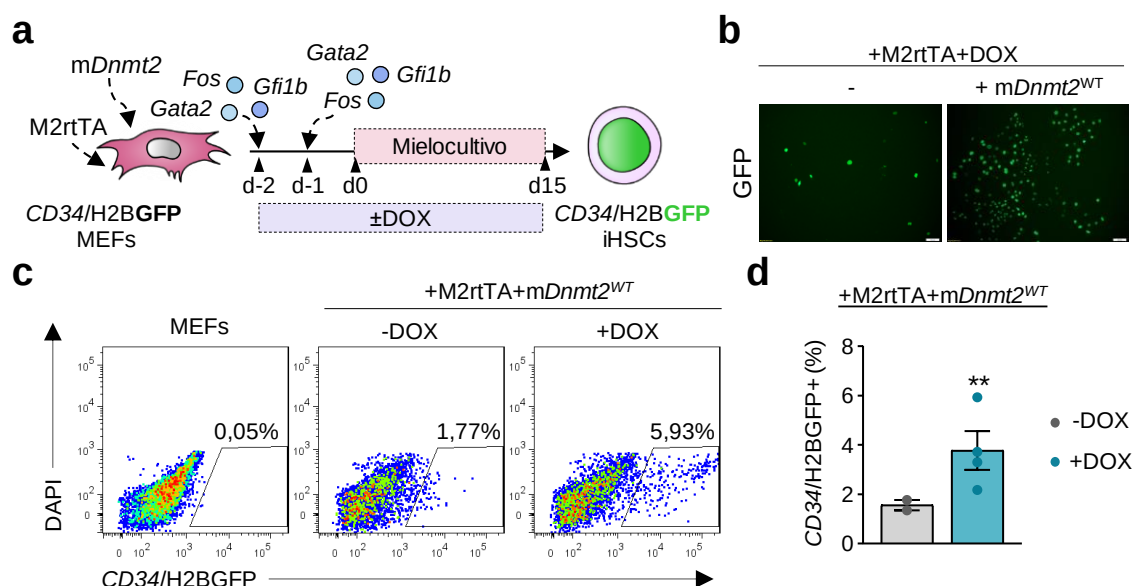


Figura R12. La sobreexpresión de *Dnmt2* facilita la obtención de precursores hematopoyéticos. **a)** Representación esquemática de la estrategia de reprogramación seguida. Los MEFs *CD34/H2BGFP* fueron transducidos con FUW-Tet-On-*mDnmt2* y FUW-M2rtTA, o solo con FUW-M2rtTA como control. Posteriormente, se indujo la sobreexpresión en dos días consecutivos de los factores de transcripción GGF y las células fueron mantenidas en medio de cultivo hematopoyético $\pm$ DOX (1 μ g/mL) durante 15 días. Luego, se evaluó la eficiencia

de la reprogramación mediante microscopía y citometría de flujo. **b)** Imágenes representativas de fluorescencia, a día 15 de reprogramación hacia iHSCs, en células transducidas con FUW-Tet-On-*mDnmt2* y FUW-M2rtTA o solo con FUW-M2rtTA, como control, en presencia de DOX. Escala: 100 μ m. **c)** Representación del porcentaje de células GFP+ obtenidas por citometría de flujo, tras la sobreexpresión de *Dnmt2* (+DOX) o en el control sin sobreexpresión (-DOX), a día 15 de reprogramación hacia iHSCs. Se muestra el porcentaje obtenido en CD34/H2BGFP MEFs parentales como control negativo. **d)** Cuantificación del porcentaje de células GFP+ obtenidas a partir de MEFs sobreexpresando *Dnmt2* (+DOX) o en el control sin sobreexpresión (-DOX), a día 15 de reprogramación. Los datos se muestran como media $\pm$ desviación estándar de cuatro réplicas biológicas. Para el cálculo de diferencias estadísticamente significativas se aplicó una prueba t de Student de dos colas no pareada y estas se indican como **: $p < 0,01$.

Finalmente, después de demostrar que el aumento de los niveles de *DNMT2* facilita una mayor eficiencia en la generación de precursores hematopoyéticos en el sistema murino, decidimos analizar si este efecto se encuentra conservado en el sistema humano y si depende de su actividad catalítica sobre el ARN. Para ello, realizamos la sobreexpresión de *hDNMT2*^{WT} o *hDNMT2*^{C79A} en fibroblastos dérmicos humanos (HDFs) mediante el sistema lentiviral inducible FUW-Tet-On (**Figura R13a**). En primer lugar, confirmamos que los niveles de expresión de ambas formas de DNMT2 en los HDFs eran similares (**Figura R13b**). Luego, se indujo la expresión de los factores de transcripción GGF en dos días consecutivos, manteniéndose las células en un medio de cultivo hematopoyético para células humanas durante 15 días (**Figura R13a**). Después, procedimos a la evaluación de la eficiencia de reprogramación mediante el conteo de colonias con morfología iHSC (**Figura R13c**). Los resultados revelaron un aumento importante en el número de colonias de iHSCs, incluso a los 10 días del proceso de reprogramación celular, cuando se sobreexpresó la forma catalíticamente activa de DNMT2 en comparación con el control (**Figura R13d**). Sin embargo, este aumento no se observó en el caso de la forma mutante catalíticamente inactiva (**Figura R13d**).

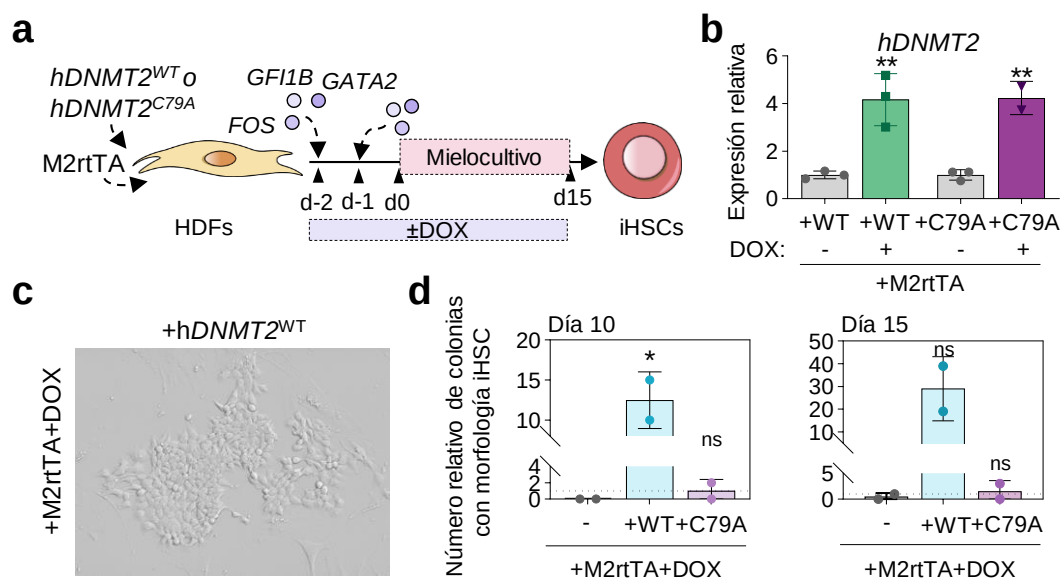


Figura R13. La ganancia de función de la actividad catalítica de DNMT2 favorece la generación de iHSCs humanas. **a)** Estrategia de reprogramación empleada. Se transdujeron HDFs con FUW-M2rtTA y FUW-Tet-On-*hDNMT2*^{WT} o FUW-Tet-On-*hDNMT2*^{C79A}, o solo con FUW-M2rtTA como control. Luego, se indujo la sobreexpresión de GGF durante dos días consecutivos y las células se mantuvieron en medio de cultivo hematopoyético durante 15 días. Se determinó el número de colonias iHSCs a los días 10 y 15. **b)** Análisis de los niveles de expresión de *hDNMT2* en HDFs transducidos con FUW-M2rtTA y FUW-Tet-On-*hDNMT2*^{WT} o FUW-Tet-On-*hDNMT2*^{C79A}, con o sin DOX, al inicio del proceso de reprogramación hacia iHSCs. Se muestran los niveles de expresión relativos a *HPRT* y al control (+M2rtTA sin DOX). **c)** Imagen representativa de una colonia de iHSCs obtenida a día 15 durante la reprogramación directa con los factores GGF y sobreexpresión de la forma catalíticamente activa de DNMT2 (*hDNMT2*^{WT}). **d)** Número de colonias iHSCs en comparación con el control (M2rtTA+DOX), a los días 10 (izquierda) y 15 (derecha) del proceso de reprogramación a iHSCs. El signo - indica el control sin sobreexpresión de *DNMT2*.

En b y d, los datos se presentan como media $\pm$ desviación estándar y el número de replicados biológicos se indica por puntos. Se utilizaron pruebas ANOVA de un factor no pareadas con corrección de Dunnett para comparaciones múltiples con el fin de calcular diferencias estadísticamente significativas, indicadas como ns: no significativo, *: $p < 0,05$ y **: $p < 0,01$.

En resumen, los resultados obtenidos hasta el momento muestran que, mientras que la pérdida de función de DNMT2 dificulta la reprogramación directa de fibroblastos hacia precursores hematopoyéticos (**Figuras R10 y R11**), su ganancia de función mejora la eficiencia de este proceso (**Figuras R12 y R13**). Además, los datos obtenidos sugieren que la actividad catalítica de DNMT2 desempeña un papel clave en los eventos de reprogramación celular directa hacia iHSCs (**Figura R13**).

6.2 DETERMINACIÓN DEL PAPEL DE DNMT2 EN LA REGULACIÓN DEL ESTADO CELULAR PLURIPOTENTE

6.2.1 La ausencia de DNMT2 es compatible con el mantenimiento de pluripotencia

Después de verificar la importancia de DNMT2 en la adquisición de pluripotencia durante la SCR, nos propusimos explorar el potencial papel de esta proteína en el mantenimiento del estado celular pluripotente.

En primer lugar, realizamos una aproximación de pérdida de función en ESCs murinas utilizando tres shRNAs independientes dirigidos contra *Dnmt2*, junto con un sh*Luci* como control. Estos shRNAs fueron subclonados en un vector lentiviral pLKO, que contiene el gen de resistencia a puomicina. Después de un periodo de 24 horas en presencia de suero y LIF (SL), procedimos a la selección de las células transducidas con cada shRNA mediante la adición de puomicina a una concentración de 2 $\mu\text{g}/\mu\text{L}$ durante dos días, con el propósito de obtener poblaciones estables que expresaran cada shRNA de interés (**Figura R14a**).

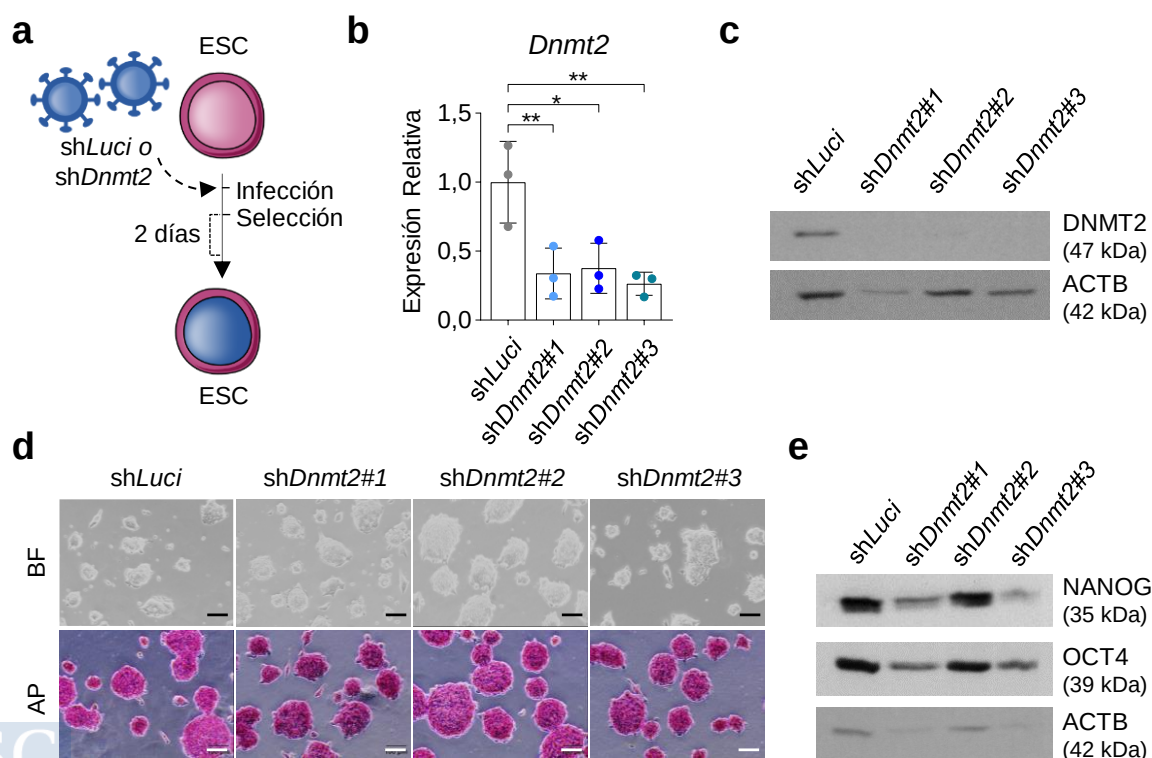


Figura R14. El silenciamiento de *Dnmt2* no altera la morfología de las ESCs. **a)** Estrategia de silenciamiento de *Dnmt2* en ESCs de ratón mediante ARN de interferencia. Las células se transdujeron con shRNAs dirigidos contra

el gen de *Luciferasa* como control (sh*Luci*) o contra *Dnmt2* (sh*Dnmt2*#1, sh*Dnmt2*#2, sh*Dnmt2*#3), en condiciones de cultivo estándar (SL). **b)** Análisis de la expresión de *Dnmt2* en ESCs transducidas con los shRNAs contra *Dnmt2* y sh*Luci* (control), mediante RT-qPCR. La expresión se muestra con relación a *Actb* y sh*Luci*. Los datos se presentan como media $\pm$ desviación estándar de triplicados técnicos. Las diferencias estadísticamente significativas se calcularon utilizando una prueba ANOVA de un solo factor no pareada, con corrección de Dunnet para comparaciones múltiples y se indican como *: $p < 0,05$ y **: $p < 0,01$. **c)** Análisis de los niveles de la proteína DNMT2 en ESCs mediante WB en las condiciones indicadas. Se muestran los niveles de ACTB como control de carga. **d)** Imágenes representativas en campo claro (BF, arriba) y tras tinción de fosfatasa alcalina (AP, abajo), de ESCs transducidas con los shRNAs indicados en condiciones estándar de cultivo. Escala: 100 μ m. **e)** Análisis mediante WB de los marcadores pluripotentes NANOG y OCT4 en ESCs transducidas con los shRNAs indicados. Se utilizaron los niveles de ACTB como control de carga.

Tras confirmar una reducción significativa en los niveles de ARN y proteína de DNMT2 con cada shRNA utilizado en comparación con el control (**Figuras R14b y R14c**), procedimos a analizar la morfología de las colonias formadas, así como su patrón de tinción de la fosfatasa alcalina (AP), en condiciones estándar de cultivo. Es importante destacar, tal y como se mencionó anteriormente, que las células pluripotentes muestran niveles elevados de AP en comparación con la mayoría de las células somáticas, por lo que la tinción de AP es ampliamente empleada como un marcador específico de pluripotencia (Stefkova et al., 2015). A continuación, observamos que después de la inhibición de *Dnmt2* las colonias de ESCs mostraban una morfología y un patrón de tinción de AP similares al grupo control (**Figura R14d**). De acuerdo con esta observación, detectamos niveles semejantes de proteína de los marcadores de pluripotencia OCT4 y NANOG, tanto en las células transducidas con cada shRNA dirigido contra *Dnmt2*, como en las transducidas con sh*Luci*, mediante WB (**Figura R14e**).

En consonancia con los resultados anteriores, no se detectaron diferencias ni en los niveles ni en la distribución de los marcadores de pluripotencia OCT4 (**Figura R15a**) y SOX2 (**Figura R15b**), mediante inmunofluorescencia, entre las ESCs transducidas con cada sh*Dnmt2* y las transducidas con sh*Luci*.

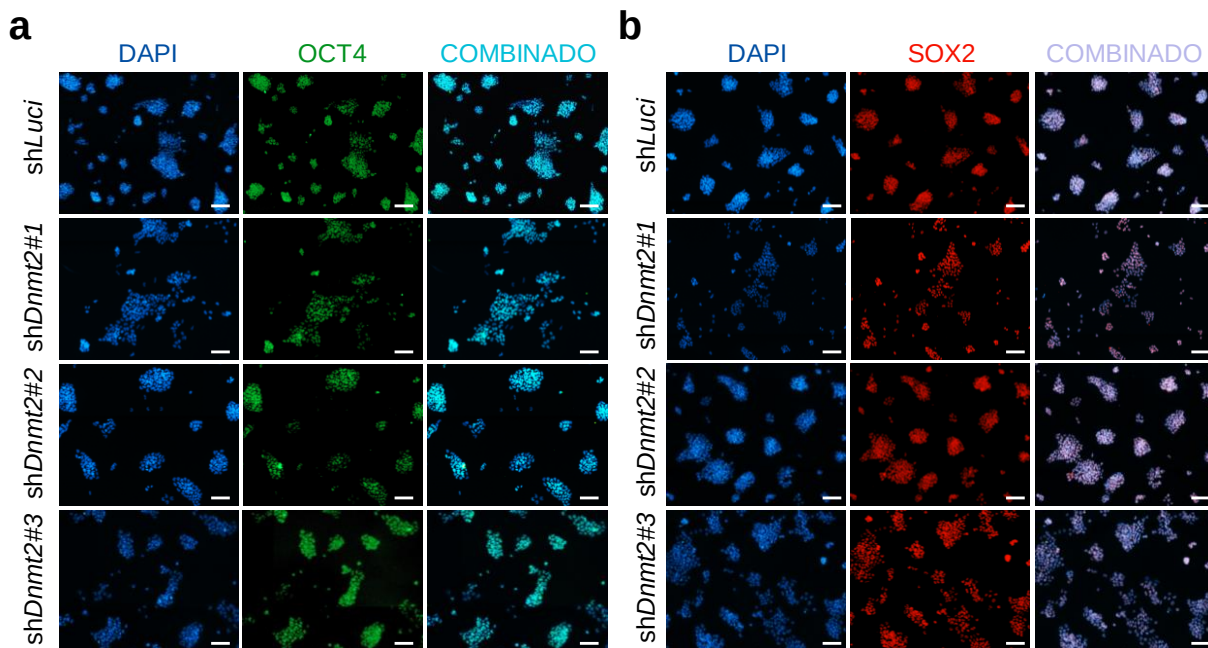


Figura R15. La inhibición de *Dnmt2* no afecta la distribución de OCT4 y SOX2 en ESCs. Imágenes de inmunofluorescencia representativas de OCT4 (a) y SOX2 (b) en ESCs transducidas con shRNAs dirigidos contra *Luciferasa* (sh*Luci*) o *Dnmt2* (sh*Dnmt2*#1, sh*Dnmt2*#2 o sh*Dnmt2*#3), en condiciones estándar de cultivo. Los núcleos fueron marcados con DAPI. Escala: 100 μ m.

A continuación, nos propusimos evaluar posibles variaciones en el ciclo celular y la proliferación de las ESCs transducidas con shRNAs dirigidos contra *Dnmt2* y *Luciferasa*. Las células pluripotentes se caracterizan por su corto ciclo celular y una alta tasa de proliferación, con una fase S más prolongada en comparación con las fases G1 y G2 (Boward et al., 2016; Orford y Scadden, 2008).

Para analizar el ciclo celular, medimos el contenido de ADN de las células mediante citometría de flujo tras su tinción con yoduro de propidio, un tinte fluorescente que se une al ADN de manera proporcional a su cantidad. Los resultados obtenidos mostraron, de conformidad con la literatura al respecto de células pluripotentes, una mayor cantidad de células en fase S en las condiciones analizadas (**Figuras R16a y R16b**). Es más, tanto las células control como las transducidas con shRNAs contra *Dnmt2* mostraron una distribución de fases del ciclo celular similar (**Figuras R16a y R16b**).

Posteriormente, con el fin de descartar posibles variaciones en la proliferación de las células pluripotentes tras la inhibición de *Dnmt2*, sembramos un número equivalente de ESCs en cada condición y las mantuvimos en medio cultivo estándar (SL) durante un período de seis días. Cada dos días, registramos el número total de células (**Figura R16c**). En este caso, tampoco observamos diferencias significativas en los patrones de proliferación celular tras la pérdida de función de DNMT2 (**Figura R16d**).

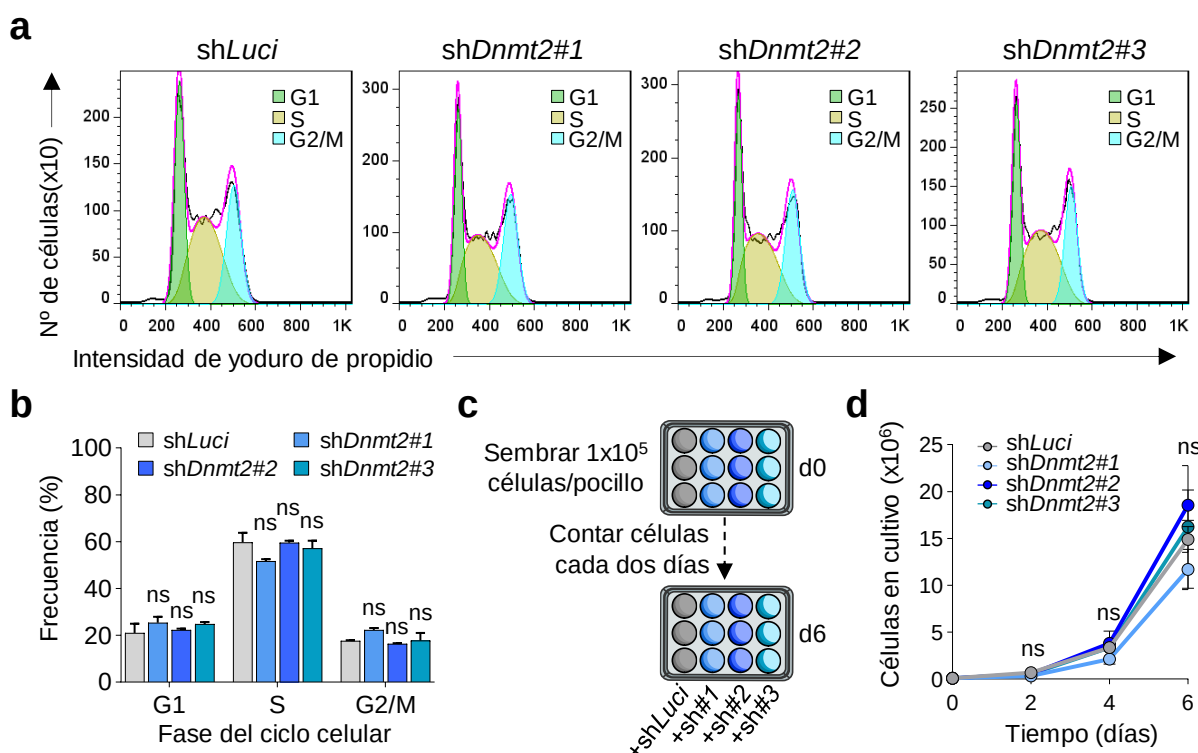


Figura R16. La pérdida de función de DNMT2 no altera la proliferación celular de las ESCs. a) Imágenes representativas de perfiles de ciclo celular obtenidos por citometría de flujo tras la tinción con yoduro de propidio en ESCs transducidas con shRNAs dirigidos contra los genes *Luciferasa* (shLuci) o *Dnmt2* (shDnmt2#1, shDnmt2#2 o shDnmt2#3), en condiciones estándar de cultivo (SL). b) Frecuencia de las diferentes fases del ciclo celular (G1, S y G2) para cada una de las condiciones analizadas. c) Esquema del análisis de proliferación en ESCs. Se sembraron 1×10^5 células por pocillo en placas de 12 pocillos en condiciones estándar de cultivo (SL) y se registró el número total de células cada dos días durante un total de seis. d) Recuento del número de células en cada condición y día correspondientes. En b y d, los datos se presentan como media $\pm$ desviación estándar de dos (b) y tres (d) réplicas biológicas. Para el cálculo de diferencias estadísticamente significativas se aplicó una prueba

ANOVA de dos factores pareada, con corrección de Dunnet para comparaciones múltiples. Estas se indican como ns: no significativo.

Finalmente, evaluamos la capacidad de autorrenovación de las ESCs después del silenciamiento de *Dnmt2* mediante un ensayo de formación de colonias (CFA), ampliamente utilizado en el campo de pluripotencia. Este método implica la siembra de ESCs a baja densidad para evaluar su comportamiento a nivel de célula individual. Las ESCs transducidas con shRNAs dirigidos contra *Dnmt2* o *Luciferasa* se mantuvieron en cultivo durante 4 días, en presencia o ausencia de LIF, una citocina que es necesaria para el mantenimiento del estado pluripotente de células de ratón en cultivo (Smith et al., 1988; Williams et al., 1988). Para evaluar la capacidad de formación de colonias pluripotentes en cada condición, realizamos una tinción de fosfatasa alcalina (AP) (**Figuras R17a y R17b**) y luego clasificamos las colonias en tres categorías: i) no diferenciadas, con tinción AP homogénea y morfología compacta; ii) parcialmente diferenciadas, con tinción AP heterogénea y morfología irregular; y iii) colonias diferenciadas, con pérdida completa de la tinción AP y morfología plana (**Figura R17c, derecha**).

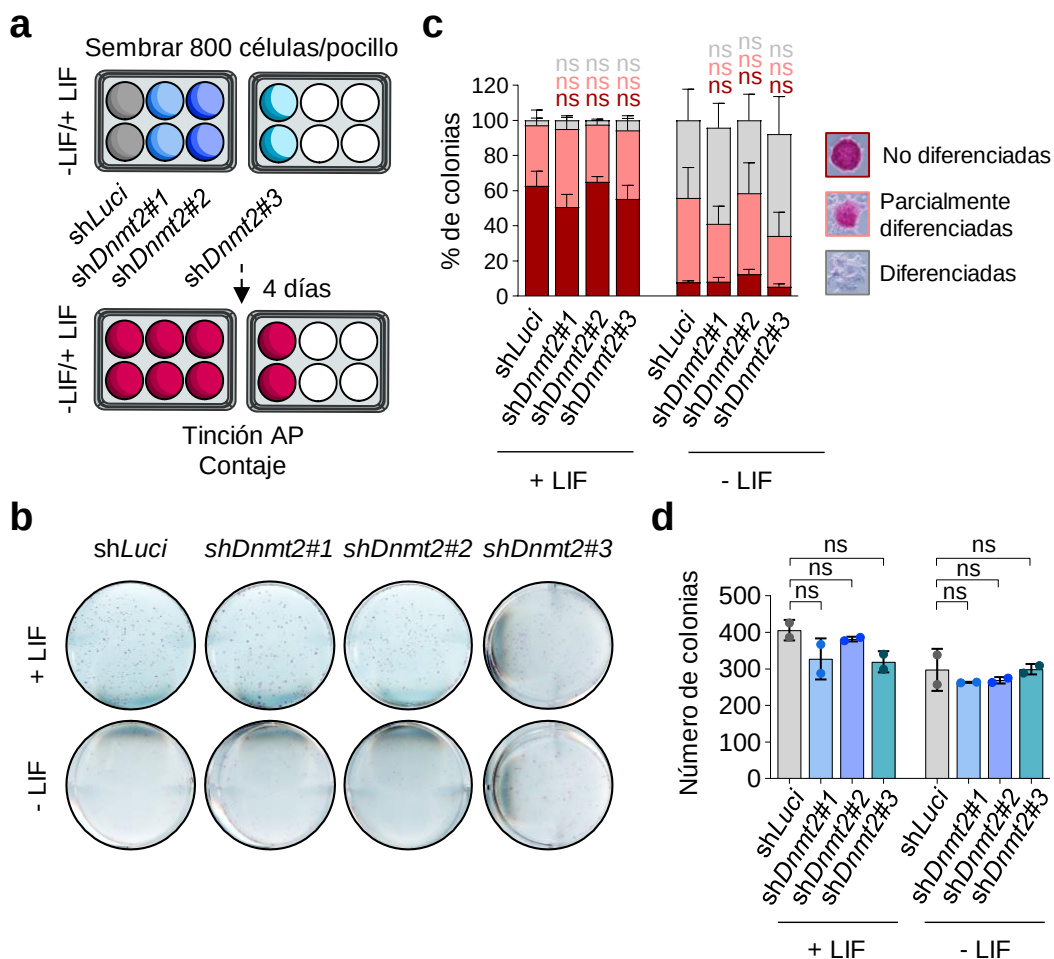


Figura R17. El silenciamiento de *Dnmt2* no afecta a la capacidad de autorrenovación de las ESCs. a) Estrategia del ensayo CFA. Se sembraron 800 células transducidas con shRNAs contra *Dnmt2* (sh*Dnmt2*#1, sh*Dnmt2*#2, sh*Dnmt2*#3) o *Luciferasa* (sh*Luci*) por pocillo en placas de 6 pocillos. Tras 4 días de cultivo, con y sin LIF, se realizó una tinción AP y se clasificaron las colonias en diferenciadas, parcialmente diferenciadas y no diferenciadas, en cada condición. b) Imágenes representativas de la tinción AP después de 4 días de cultivo en presencia y ausencia de LIF, en cada condición indicada. c) Recuento del número de colonias diferenciadas, parcialmente diferenciadas y no diferenciadas en cada condición indicada. d) Recuento del número total de colonias generadas en el ensayo de formación de colonias en cada una de las condiciones presentadas. En c y d,

los datos se presentan como media $\pm$ desviación estándar de dos réplicas biológicas. Para el cálculo de diferencias estadísticamente significativas se aplicó una prueba ANOVA de dos factores pareada (c), o de un factor no pareada (d), con corrección de Dunnet para comparaciones múltiples. Estas se indican como ns: no significativo.

Es importante destacar que, en condiciones estándar de cultivo, las ESCs pueden mostrar una variedad de estados pluripotentes alternativos, lo que puede resultar en la diferenciación espontánea de algunas células y la formación de colonias parcial o completamente diferenciadas incluso en presencia de LIF (Guo et al., 2016; Hayashi et al., 2008). Por otro lado, la falta de LIF promueve la salida del estado pluripotente y la consecuente pérdida de tinción AP (Guo et al., 2016; Hayashi et al., 2008).

Los resultados obtenidos no mostraron diferencias significativas entre las colonias control (sh*Luci*) y las derivadas de células transducidas con sh*Dnmt2*, ni en presencia ni en ausencia de LIF, en ninguna de las categorías definidas (**Figuras R17b y R17c**). Además, tampoco se detectaron diferencias en el número total de colonias entre las condiciones estudiadas (**Figuras R17d**).

En conjunto, los resultados obtenidos muestran que la reducción de los niveles de expresión de *Dnmt2* es compatible con el mantenimiento de un estado celular pluripotente en condiciones estándar de cultivo.

Posteriormente, para confirmar que la función de DNMT2 no es esencial en el mantenimiento del estado celular pluripotente y considerando que la reducción de su expresión obtenida mediante interferencia de ARN podría enmascarar el efecto de su ausencia total, optamos por generar una línea de ESCs *Dnmt2* knockout (KO) utilizando la tecnología CRISPR/Cas9 (Jinek et al., 2012). En resumen, diseñamos guías de ARN específicas para dirigir la endonucleasa Cas9 al primer exón del gen *Dnmt2* (**Figura R18a**).

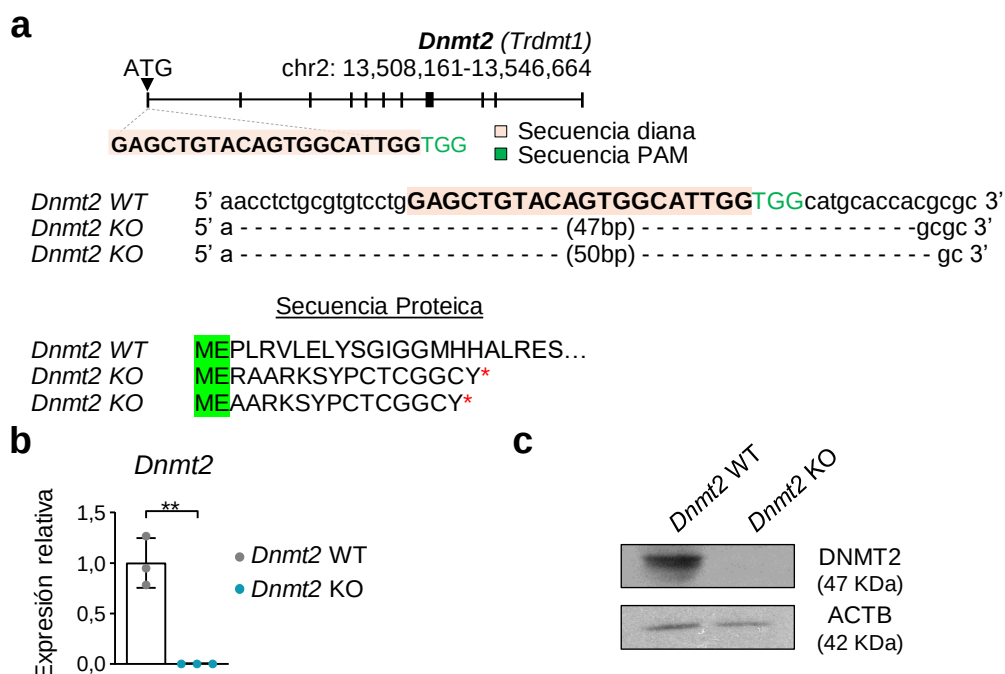


Figura R18. Generación de una línea de ESCs KO para *Dnmt2*. a) Estrategia empleada para generar una línea celular en ESCs con la delección del gen *Dnmt2* utilizando el sistema CRISPR/Cas9. La flecha señala el codón de inicio de la traducción (ATG), junto con el exón 1 y la secuencia diana resaltada en naranja. La secuenciación Sanger confirmó la mutación generada, que implicó la eliminación de 47 pares de bases (pb) en un alelo y 50 pb en el otro. La simulación de la secuencia proteica resultante (panel inferior) revela la presencia de sitios de terminación prematura de la traducción (asterisco). b) Análisis de los niveles de expresión de *Dnmt2* mediante

RT-qPCR en las ESCs WT y KO para *Dnmt2*. La expresión se muestra con relación a *Hprt* y *shLuci*. Los datos se presentan como media $\pm$ desviación estándar de triplicados biológicos. Para analizar las diferencias estadísticas, se realizó una prueba t de Student de una cola no pareada, donde **: $p < 0,01$. c) Análisis de los niveles de proteína de DNMT2 mediante WB. Se empleó ACTB como control de carga.

A continuación, se procedió a la transfección de ESCs WT con un plásmido que albergaba la endonucleasa, la guía de ARN y un gen de resistencia al antibiótico puromicina, que permitió la selección transitoria de las células que incorporaron el vector. Después, obtuvimos clones independientes cuya homocigosidad para la delección de *Dnmt2* se evaluó mediante una estrategia de genotipado por PCR (datos no mostrados). Se extrajo ADN de las poblaciones clonales candidatas para analizar las delecciones mediante secuenciación Sanger y en base a los resultados, elegimos un clon de ESCs con una delección de 47 pares de bases en un alelo y 50 pares de bases en el otro (**Figura R18a**), ambas generando proteínas con sitios de parada de la traducción prematuros. La ausencia de expresión del gen *Dnmt2* fue validada por RT-qPCR (**Figura R18b**) y la pérdida completa de los niveles de proteína, mediante WB (**Figura R18c**).

Tras la generación de la línea de ESCs *Dnmt2* KO, evaluamos su capacidad para formar colonias pluripotentes, así como posibles defectos en la expresión de marcadores de pluripotencia. En primer lugar, valoramos la morfología de las colonias formadas por células WT y KO para *Dnmt2*, así como los niveles del marcador AP (**Figura R19a**). De conformidad con los resultados anteriores, utilizando shRNAs contra este gen, no se observaron diferencias evidentes ni en la morfología de las colonias generadas, ni en los patrones de tinción AP (**Figura R19a**). Además, tampoco se detectaron cambios sustanciales en los niveles de proteína de los marcadores de pluripotencia OCT4 y NANOG (**Figura R19b**).

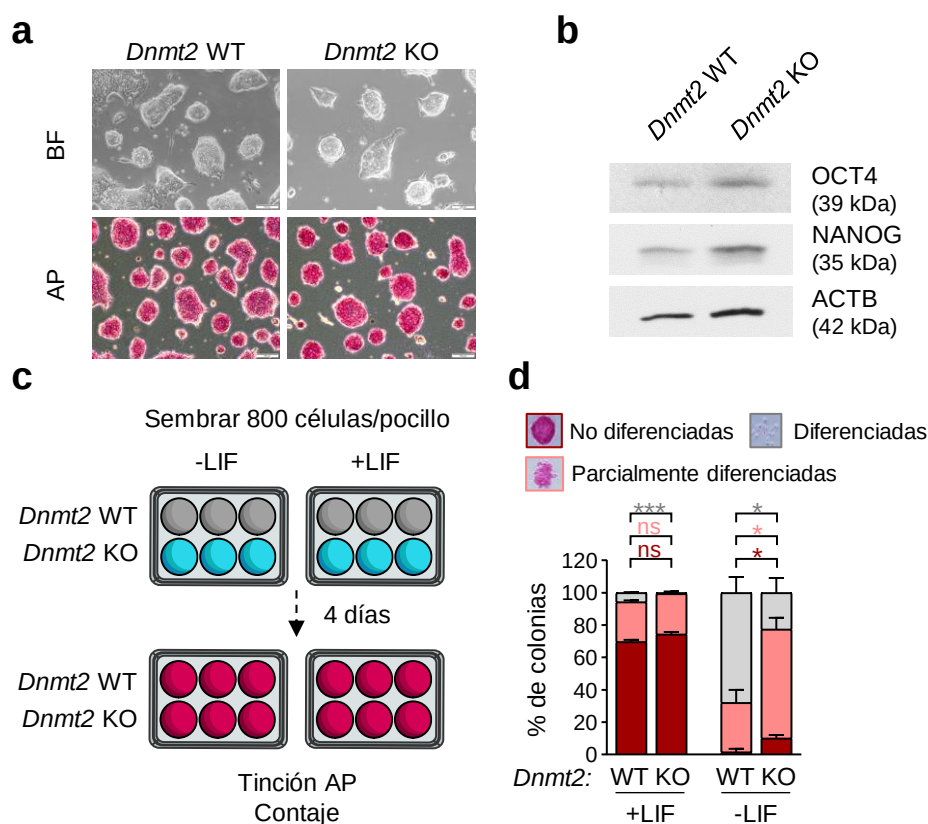


Figura R19. La ausencia de DNMT2 no compromete la capacidad de autorrenovación de las células pluripotentes. a) Imágenes representativas en campo claro (BF, panel superior) y tinción de fosfatasa alcalina

(AP, panel inferior), de ESCs WT o KO para *Dnmt2* en condiciones estándar de cultivo. Escala: 100 μ m. **b)** Análisis mediante WB de los niveles de proteína de los marcadores pluripotentés NANOG y OCT4 en células WT y KO para *Dnmt2*. Se muestran los niveles de ACTB como control de carga. **c)** Esquema de la estrategia seguida en el ensayo de formación de colonias pluripotentes. Se sembraron 800 células WT o KO para *Dnmt2* por pocillo en placas de 6 pocillos y tras 4 días en cultivo en presencia y ausencia de LIF, se realizó una tinción AP y se clasificaron las colonias en: diferenciadas, parcialmente diferenciadas y no diferenciadas, para cada condición. **d)** Recuento del número de colonias y clasificación de estas en los grupos previamente establecidos. Los datos se expresan como media $\pm$ desviación estándar de triplicados biológicos. Para el cálculo de diferencias estadísticamente significativas se aplicó una prueba ANOVA de dos factores pareada con corrección de Tukey para comparaciones múltiples. Se indican como ns: no significativo, *: $p < 0,05$ y ***: $p < 0,001$.

Posteriormente, llevamos a cabo un ensayo de formación de colonias para evaluar la capacidad de autorrenovación de las ESCs en ausencia de DNMT2 (**Figura R19c**). En este caso, se observó una reducción significativa en el número de colonias diferenciadas después de la pérdida total de DNMT2, tanto en presencia como en ausencia de LIF (**Figura R19d**). Además, en ausencia de LIF se detectó un aumento significativo en el número de colonias no diferenciadas (**Figura R19d**), lo que sugiere una retención de características pluripotentes en las ESCs KO para *Dnmt2*, no observada previamente en las ESCs transducidas con shRNAs dirigidos contra este gen (**Figura R17c**).

Finalmente, se analizó el patrón del ciclo celular y la tasa de proliferación de las ESCs *Dnmt2* KO en comparación con las ESCs WT. De acuerdo con los resultados previos, no se observaron diferencias significativas ni entre los patrones de ciclo celular (**Figuras R20a y R20b**), ni entre la tasa de proliferación (**Figuras R20c y R20d**) de células *Dnmt2* WT y *Dnmt2* KO, en condiciones estándar de cultivo.

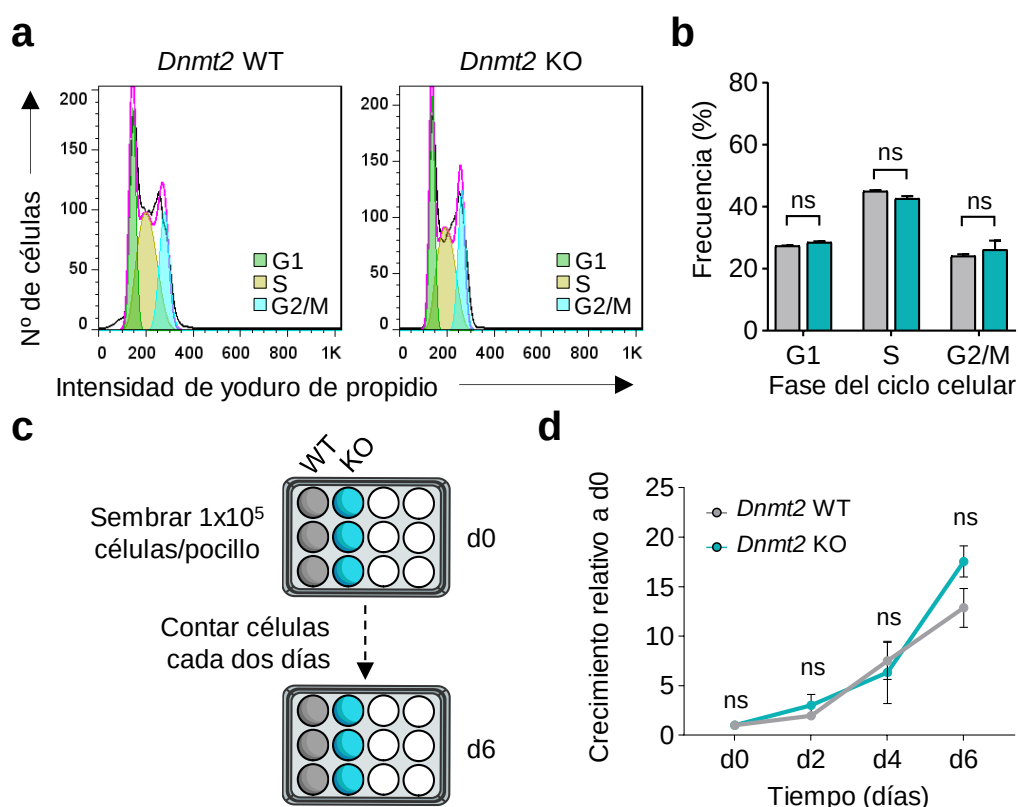


Figura R20. La pérdida de DNMT2 no afecta a la proliferación de las ESCs. **a)** Imágenes representativas de los perfiles del ciclo celular en ESCs WT y KO para *Dnmt2*, obtenidos por citometría de flujo tras la tinción con yoduro de propidio, en condiciones estándar de cultivo. **b)** Cuantificación del análisis de ciclo celular. Se representa el porcentaje de células en las distintas fases del ciclo celular (G1, S y G2), para las dos condiciones

analizadas. c) Representación esquemática del método seguido para el análisis de proliferación celular en ESCs WT o KO para *Dnmt2*. Se sembraron 1×10^5 células por pocillo en placas de 12 pocillos para cada una de las condiciones analizadas en condiciones estándar de cultivo y se contó el número total de células cada dos días durante un total de seis. d) Representación del crecimiento relativo respecto a día 0 de ESCs WT o KO para *Dnmt2* en los días especificados. En b y d, los datos se presentan como media $\pm$ desviación estándar de dos (b) y tres (d) réplicas biológicas. Para el cálculo de diferencias estadísticamente significativas se aplicaron pruebas t de Student entre las condiciones analizadas, con una corrección FDR para comparaciones múltiples, ns: no significativo.

En resumen, los resultados obtenidos, tanto después del silenciamiento como de la eliminación de *Dnmt2*, indican que su función no es esencial para el mantenimiento del estado celular pluripotente. Sin embargo, la significativa retención de características pluripotentes en las ESCs KO para *Dnmt2* en ausencia de LIF sugiere que podría desempeñar un papel relevante en eventos de diferenciación temprana de las células pluripotentes en cultivo.

6.2.2 La ausencia de DNMT2 altera el transcriptoma de las ESCs

Previamente, hemos demostrado que la función de DNMT2 no es esencial para el mantenimiento del estado celular pluripotente. Sin embargo, su ausencia parece favorecer la capacidad de autorrenovación de las ESCs en ausencia de LIF (**Figura R19d**), lo que sugiere un potencial papel de esta proteína en la especificación celular temprana durante la salida de pluripotencia. Teniendo esto en cuenta, nos propusimos investigar los posibles programas moleculares en los que participa DNMT2 en el estado celular pluripotente. Para abordar este objetivo, llevamos a cabo un análisis de transcriptoma completo de las ESCs WT y KO para *Dnmt2*, cultivadas en condiciones estándar de cultivo (SL), mediante RNA-seq (**Figura R21a**). Con estos datos y aplicando el coeficiente de correlación de Pearson pudimos generar una matriz de similitud acoplada a una agrupación jerárquica, que nos permitió clasificar diferencialmente los transcriptomas de las ESCs WT y KO para *Dnmt2* (**Figura R21b**). Asimismo, mediante un análisis de componentes principales (PCA) observamos que la ausencia de DNMT2 causó alteraciones transcripcionales significativas, suficientes para establecer dos grupos distintos: WT y KO (**Figura R21c**).

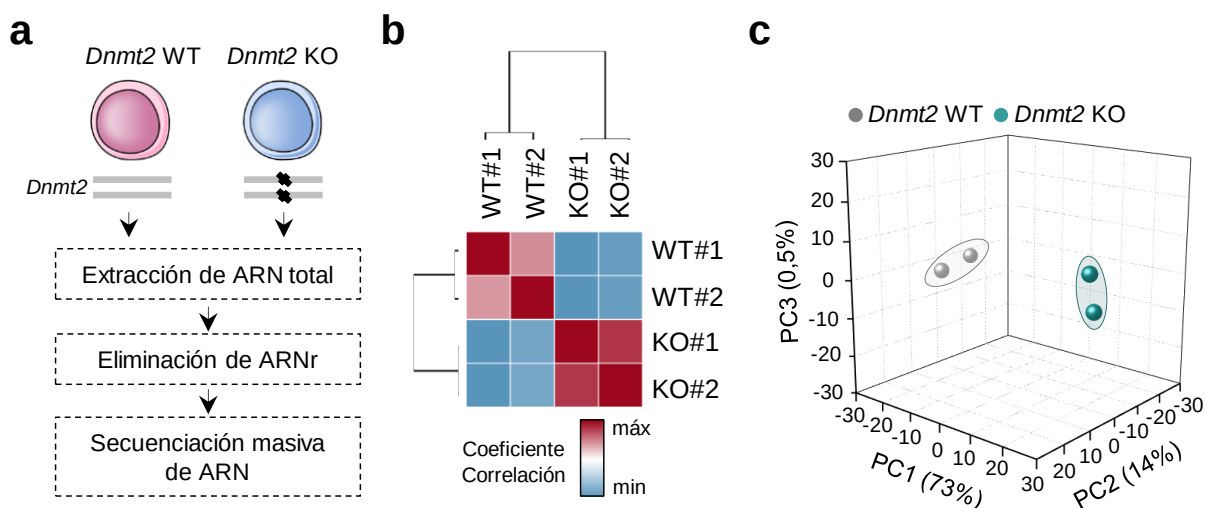


Figura R21. Análisis de los transcriptomas de las ESCs WT y KO para *Dnmt2*. a) Esquema de la estrategia seguida para el análisis del transcriptoma de ESCs WT y KO para *Dnmt2*. Se emplearon duplicados biológicos de cada una de las condiciones estudiadas, que fueron mantenidos en condiciones estándar de cultivo. Posteriormente, se procedió a la extracción de ARN total, la eliminación del ARNr y la preparación de librerías para la secuenciación masiva de ARN (RNA-seq). b) Análisis de los perfiles transcripcionales de ESC WT y KO para *Dnmt2* mediante una matriz de similitud acoplada a una agrupación jerárquica en base al coeficiente de

correlación de Pearson. c) PCA de los perfiles transcripcionales de las células ESCs WT y KO para *Dnmt2*. El porcentaje de variación explicado por cada componente principal (PC) se indica en el eje correspondiente.

A continuación, llevamos a cabo un análisis completo de los transcriptomas de las ESCs WT y KO para *Dnmt2* con el objetivo de explorar los posibles programas moleculares en los que DNMT2 podría estar involucrado en células pluripotentes. En primer lugar, identificamos los transcritos cuya expresión variaba de manera significativa entre las condiciones analizadas, con una tasa de descubrimientos falsos inferior al 5% ($FDR < 0,05$). Bajo este criterio, detectamos que la pérdida de DNMT2 ocasiona una desregulación significativa en la expresión de 2.117 transcritos (**Figura R22a**), siendo la mayoría de los afectados (>90%) codificantes de proteínas (**Figura R22b**).

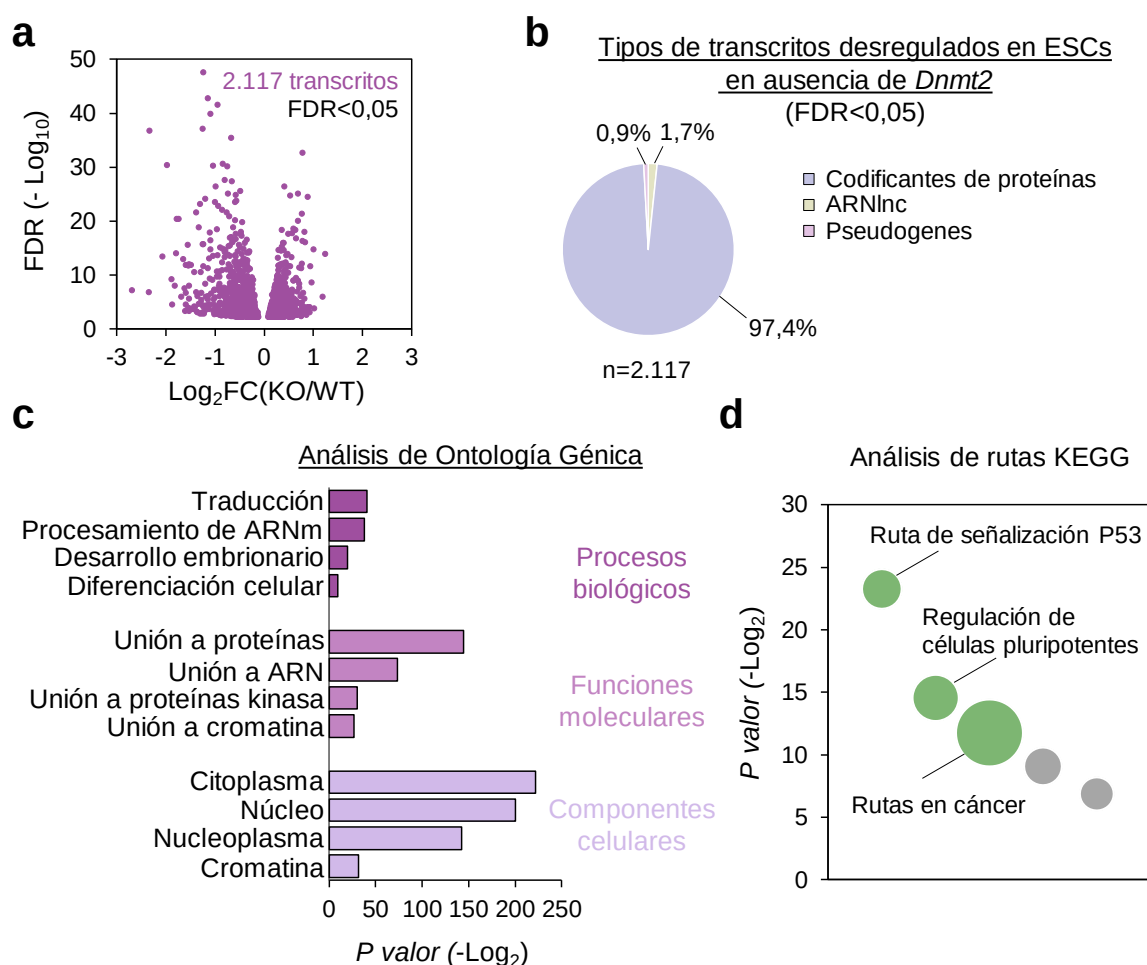


Figura R22. La falta de DNMT2 altera el transcriptoma de las ESCs. a) Diagrama de volcán que muestra los transcritos diferencialmente expresados ($FDR < 0,05$) en ausencia de DNMT2 en ESCs. b) Proporción de los tipos de transcritos diferencialmente expresados en ESCs tras la pérdida de DNMT2. Se indican: transcritos codificantes de proteínas, transcritos largos no codificantes (ARNlnc) y pseudogenes. c) Análisis de ontología genética de los genes diferencialmente expresados en ausencia de DNMT2. Se muestran procesos biológicos, funciones moleculares y componentes celulares con los que dichos transcritos están relacionados. d) Análisis de rutas KEGG en las que están implicados los transcritos diferencialmente expresados tras la eliminación de DNMT2 en ESCs.

Con el fin de comprender las implicaciones biológicas de la ausencia de DNMT2 en ESCs, llevamos a cabo un análisis de ontología genética (GO) que nos permitió identificar posibles procesos y vías en los que esta proteína podría estar involucrada. Para realizar este análisis, empleamos la herramienta en línea DAVID (Huang da et al., 2009; Sherman et al.,

2022). Este análisis mostró un enriquecimiento de los genes diferencialmente expresados en ausencia de DNMT2 (FDR<0,05) en procesos relacionados con la pluripotencia, incluyendo diferenciación celular y desarrollo embrionario uterino, entre otros (**Figura R22c**).

El posterior análisis de rutas KEGG, realizado también a través de la herramienta DAVID, respaldó estas observaciones al mostrar que algunos de los genes desregulados estaban asociados con vías de señalización implicadas en la regulación de la pluripotencia (**Figura R22d**).

Además, el análisis de rutas KEGG reveló que parte de los transcritos significativamente desregulados tras la eliminación de *Dnmt2* estaban implicados en la vía de señalización de la proteína P53, así como en otras relacionadas con el cáncer (**Figura R22d**). Para validar esta observación, analizamos los niveles de la proteína P53 en ESCs *Dnmt2* WT y KO mediante WB, observándose una reducción de estos después de la eliminación de DNMT2 (**Figura R23a**). Sin embargo, los niveles de expresión del gen *Trp53*, que codifica esta proteína, no varían significativamente en las ESCs *Dnmt2* KO según los resultados de RNA-seq (**Figura R23b**) y la posterior validación por RT-qPCR (**Figura R23c**). Esto sugiere un cambio en la regulación de P53 a nivel postranscripcional y/o de traducción, en ausencia de DNMT2.

Considerando esto, empleamos los datos generados por RNA-seq para analizar los niveles de expresión del principal regulador negativo de P53, MDM2, que regula la actividad de esta proteína a través de diversos mecanismos, como la degradación proteasomal, la exportación nuclear y el bloqueo de su transactivación, entre otros (Hafner et al., 2019; Levine, 2020). En este caso, observamos una regulación al alza significativa en los niveles de ARN del gen *Mdm2* tras la pérdida de DNMT2 (**Figura R23d**), que fue validada por RT-qPCR (**Figura R23e**). Esto sugiere la existencia de un potencial mecanismo de regulación positiva de DNMT2 sobre P53 a través de la represión transcripcional del gen *Mdm2*.

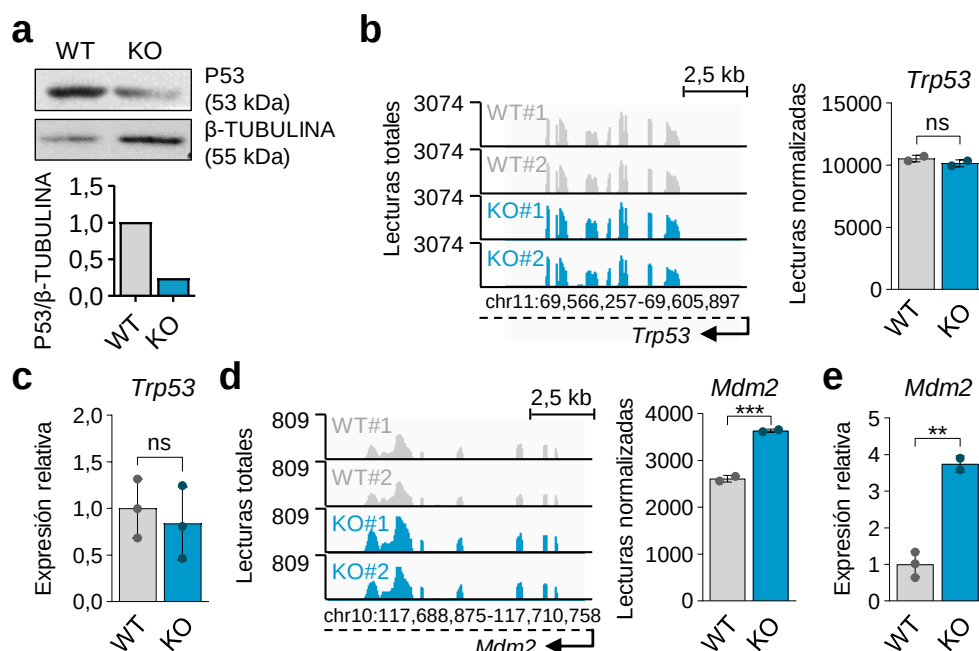


Figura R23. Efecto de la pérdida de DNMT2 en los niveles de la proteína P53 en ESCs. a) Niveles de la proteína P53 en ESCs WT y KO para *Dnmt2* evaluados por WB (panel superior). Se utilizó β-TUBULINA como control de carga. En el panel inferior se muestran los niveles de P53 relativizados a β-TUBULINA para cada condición. b) Visualización integral de los niveles globales de expresión del gen *Trp53* en las muestras analizadas mediante RNA-seq (izquierda), y niveles de expresión de este gen obtenidos por RNA-seq, representados como lecturas

normalizadas (derecha). c) Expresión relativa del gen *Trp53* obtenida mediante RT-qPCR en ESCs con y sin *Dnmt2*. Se muestra los niveles relativos a los de *Hprt* y el control (WT). d) Visualización integral de los niveles globales de expresión del gen *Mdm2* en las muestras analizadas mediante RNA-seq (izquierda) y niveles de expresión de este gen obtenidos por RNA-seq, representados como lecturas normalizadas (derecha). e) Análisis de los niveles de expresión del gen *Mdm2* por RT-qPCR en ESCs *Dnmt2* WT y KO. Se muestra los niveles relativos a los de *U6* y el control (WT). En b-e, los datos se presentan como media $\pm$ desviación estándar y el número de replicados biológicos se indica por puntos. Las diferencias estadísticamente significativas se determinaron mediante una prueba t de Student de una cola no pareada, ns: no significativo, **: $p < 0,05$, ***: $p < 0,001$.

Dado que P53 desempeña un papel fundamental en el control de la división celular y la eliminación de células dañadas o con anomalías genéticas, y que mutaciones y cambios en los niveles de este gen han sido ampliamente documentados en muchos tipos de cáncer (Muller y Vousden, 2014; Sabapathy y Lane, 2018), estos resultados en ESCs murinas nos motivaron a abrir una nueva línea de investigación centrada en dilucidar el posible papel de DNMT2 en el contexto del cáncer, que se detallará en la sección 6.3 de este manuscrito.

Nuestros resultados muestran que la ausencia de DNMT2 conlleva un aumento en el número de colonias pluripotentes en condiciones de diferenciación (sin LIF) (**Figura R19d**). Asimismo, los análisis *in silico* revelan que los genes significativamente desregulados ($FDR < 0,05$) tras la pérdida de DNMT2 en ESCs, se encuentran implicados en procesos de diferenciación y desarrollo, así como en vías moleculares vinculadas a la regulación de la pluripotencia (**Figuras R22c y R22d**).

Teniendo esto en cuenta, y que una de las principales características de las células pluripotentes es su capacidad de diferenciación a todos los tipos celulares que componen un organismo adulto, planteamos la posibilidad de que DNMT2 intervenga en la regulación de la expresión de genes implicados en procesos de diferenciación temprana. En este contexto, esperaríamos que los niveles de expresión de marcadores de diferenciación temprana disminuyeran en las ESCs KO para *Dnmt2*. Para validar esta hipótesis, clasificamos los genes diferencialmente expresados en ausencia de DNMT2 en dos categorías: i) genes cuya expresión disminuye y ii) genes cuya expresión aumenta. Además, aplicamos un criterio de cambio en los niveles de expresión de al menos el 50% ($FC > 1,5$), lo que nos permitió seleccionar los genes con los cambios más significativos entre ESCs *Dnmt2* KO y WT. Siguiendo esta estrategia, identificamos una regulación al alza en 70 transcritos y una regulación a la baja en 265 transcritos (**Figura R24a**). Dado el contexto de nuestra hipótesis, nos centramos en los transcritos cuya expresión disminuía en ausencia de DNMT2.

A continuación, procedimos a determinar en qué procesos biológicos se encontraban involucrados los transcritos cuya expresión disminuía después de la eliminación de DNMT2. Para este propósito, empleamos el sistema de clasificación en línea PANTHER (Mi et al., 2019; Thomas et al., 2022). En concordancia con nuestra hipótesis inicial, los resultados de este análisis revelaron que los genes cuya expresión se ve disminuida tras la pérdida de DNMT2 están relacionados con procesos de diferenciación, como la morfogénesis en la diferenciación y la diferenciación celular (**Figura R24b**). Además, se identificó la participación de estos transcritos en el desarrollo epitelial (**Figura R24b**), lo cual coincide con nuestras observaciones previas durante la adquisición de pluripotencia, que mostraban un defecto en el proceso de transición mesenquimal-epitelial (MET) en ausencia de DNMT2 (secciones 6.1.2 y 6.1.3).

Centrándonos en el posible papel de DNMT2 durante la diferenciación celular, realizamos un análisis de enriquecimiento génico (GSEA) utilizando una lista predefinida de genes involucrados en este proceso (GO:0030154). Los resultados de este análisis fueron coherentes con lo observado en el análisis de enriquecimiento en procesos biológicos (**Figura**

R24b), indicando que los niveles de expresión del transcriptoma de las ESCs WT están enriquecidos de manera significativa en genes relacionados con la diferenciación celular en comparación con las ESCs KO para *Dnmt2* (**Figura R24c**).

Para respaldar estas observaciones, elegimos cuatro de los genes principales cuya expresión disminuía notablemente en ausencia de DNMT2 ($FDR < 0,05$, $FC > 1,5$) y que desempeñan un papel crucial en la diferenciación celular (*Hexa*, *Map1b*, *Smyd1* y *Wnt7b*) (**Figura R24d**). Posteriormente, confirmamos esta disminución en sus niveles de expresión mediante RT-qPCR (**Figura R24e**).

Colectivamente, estos resultados respaldan la noción de que DNMT2 desempeña un papel en la regulación transcripcional de genes vinculados al proceso de diferenciación celular, favoreciendo su expresión.

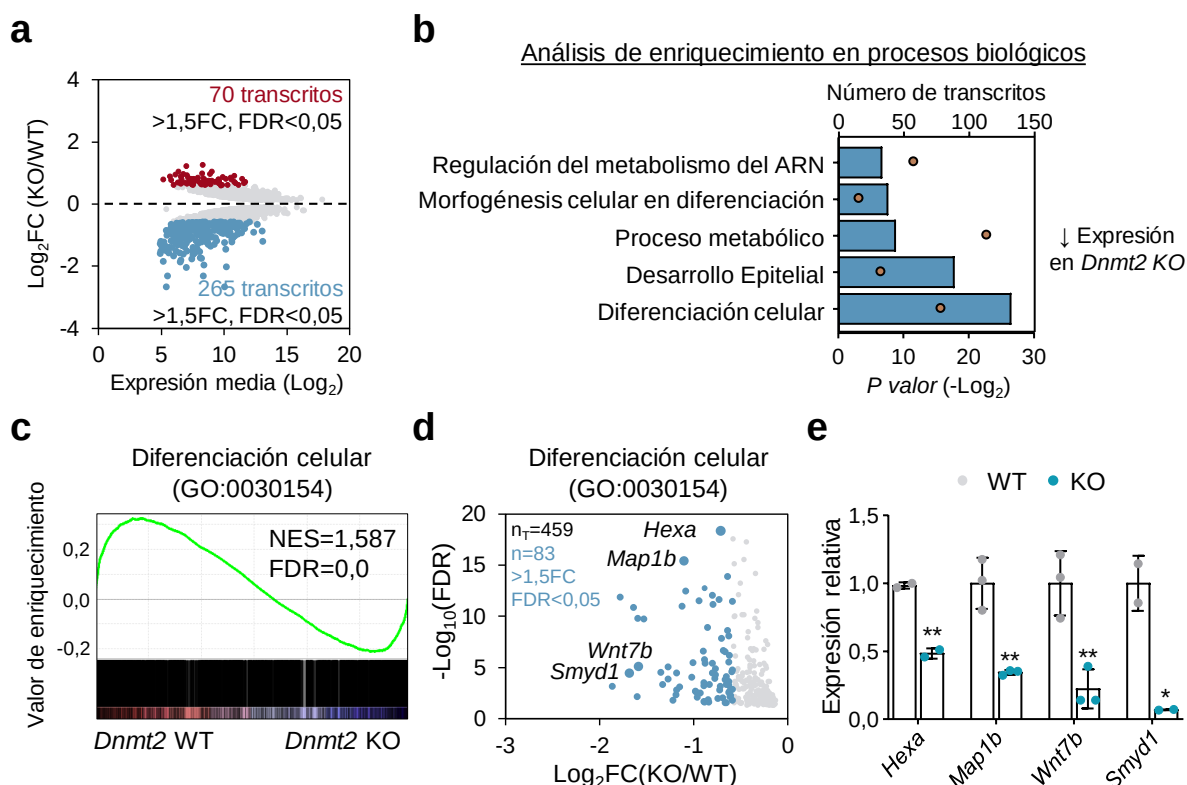


Figura R24. La ausencia de DNMT2 afecta a la correcta activación de genes implicados en la diferenciación celular en ESCs. a) Representación de los transcritos regulados diferencialmente al alza (rojo) o a la baja (azul) en ESCs KO para *Dnmt2* en comparación con ESCs WT ($FDR < 0,05$ y $FC > 1,5$). b) Análisis de los procesos biológicos afectados por los transcritos significativamente inhibidos en ausencia de DNMT2. Los valores de p ($-Log_2$) se representan mediante barras, y el número de genes en cada categoría se indica mediante puntos. c) Análisis de enriquecimiento de conjunto de genes (GSEA) de los datos de RNA-seq de ESCs WT y KO para *Dnmt2*, frente a una firma génica de diferenciación (GO:0030154). Se indican la puntuación de enriquecimiento normalizada (NES) y el FDR. d) Representación de los niveles de expresión de 459 genes implicados en diferenciación (n_T), señalando en azul aquellos significativamente regulados a la baja tras la pérdida de DNMT2 ($n=83$). e) Validación de los niveles de expresión de los genes indicados mediante RT-qPCR. La expresión se muestra con relación a *Hprt* y a ESCs WT. Los datos se presentan como media $\pm$ desviación estándar de triplicados técnicos. Se utilizó una prueba t de Student de una cola no pareada para calcular diferencias estadísticamente significativas, indicadas como * ($p < 0,05$) y ** ($p < 0,01$).

Por otro lado, es importante destacar que estudios previos han revelado que la pérdida de función de DNMT2 tiene un impacto negativo en el silenciamiento de ciertos elementos transponibles (TEs) en *Drosophila melanogaster* (Genenncher et al., 2018; Phalke et al., 2009).

Estos elementos móviles del genoma constan de dos clases principales: los transposones de ADN, cuyo desplazamiento en el genoma se produce mediante la acción de una enzima transposasa generalmente codificada por la secuencia del propio elemento; y los retrotransposones, cuya movilidad en el genoma depende de un intermediario de ARN generado a través de la transcripción del elemento (**Figura R25a**) (Deniz et al., 2019; Kazazian, 2004).

En la mayoría de los mamíferos, los transposones de ADN se encuentran inactivos (Feschotte y Pritham, 2007; Rodríguez-Terrones y Torres-Padilla, 2018). Teniendo esto en cuenta y que evidencias sugieren que los patrones de expresión de retrotransposones desempeñan un papel en la regulación del estado celular pluripotente (Fu et al., 2021; Macfarlan et al., 2012), lo que podría ser relevante en procesos de especificación temprana en la diferenciación (He et al., 2021b; Pontis et al., 2022), nos propusimos examinar las alteraciones de los retrotransposones que ocurren a nivel transcripcional tras la pérdida de DNMT2 en ESCs.

Los retrotransposones se pueden subdividir en diferentes subclases, superfamilias, familias, subfamilias y elementos (**Figura R25a**) (Deniz et al., 2019; Rodríguez-Terrones y Torres-Padilla, 2018). Para determinar si la eliminación de DNMT2 provocaba cambios en su expresión, utilizamos los datos de RNA-seq de las ESCs *Dnmt2* WT y KO para analizar los niveles de los elementos pertenecientes a familias específicas de retrotransposones (**Figura R25a**). Los resultados obtenidos revelaron la ausencia de diferencias evidentes entre los niveles de expresión de las familias de retrotransposones estudiadas en células *Dnmt2* KO en comparación con las WT (**Figura R25b**). Sin embargo, se observó una tendencia a la reducción en los niveles de expresión de la familia de retrovirus endógenos (ERV), específicamente de ERV1 (**Figura R25b**).

Puesto que cada familia está constituida por un elevado número de elementos (**Figura R25b**), lo que podría enmascarar el efecto de la ausencia de DNMT2 sobre la expresión de un grupo más reducido de estos, decidimos llevar a cabo un análisis más específico para examinar directamente qué subfamilias se encontraban significativamente desreguladas ($FDR < 0,05$) en ESCs KO para *Dnmt2*, en comparación con las WT (**Figura R25c**). El análisis reveló que 19 subfamilias de retrotransposones, pertenecientes a las familias B1, ERVK, ERV1 y L1, experimentaron cambios significativos en su expresión después de la pérdida de DNMT2 (**Figura R25c**). De estas subfamilias, 9 mostraron un aumento en su expresión (por ejemplo, RLTR9E, RLTR19-int, ETnERV-int, MMETn-int o MMERVK10C-int), mientras que otras 10 presentaron una disminución en su expresión (por ejemplo, RMER13B, L1MdTf_I/II, RLTR16B_MM o MLTR31F_MM) (**Figuras R25c y R25d**).

Es importante señalar que los mecanismos de regulación de estos elementos son altamente precisos, especialmente en las etapas tempranas del desarrollo, donde desempeñan papeles fundamentales en el control de la expresión (Chuong et al., 2013; Sakashita et al., 2023). En consonancia con esto, algunos retrotransposones desempeñan un papel clave en la regulación del estado celular pluripotente, siendo su función necesaria para el silenciamiento de marcadores de pluripotencia durante eventos de diferenciación (Morales-Hernández et al., 2016). Además, los patrones de expresión de familias específicas de retrotransposones (por ejemplo, L1 y ERVK) se han asociado con diferentes estados alternativos de pluripotencia (Fu et al., 2021), lo que sugiere que la alteración de su expresión puede causar defectos en el potencial de diferenciación de las células pluripotentes.



En resumen, estos hallazgos respaldan la hipótesis que hemos planteado, evidenciando que la ausencia de DNMT2 conduce a una perturbación en los programas transcripcionales,

afectando tanto a genes (**Figura R24**) como a TEs (**Figura R25**), lo cual podría estar vinculado a posibles defectos en la capacidad de diferenciación de las ESCs KO para *Dnmt2*.

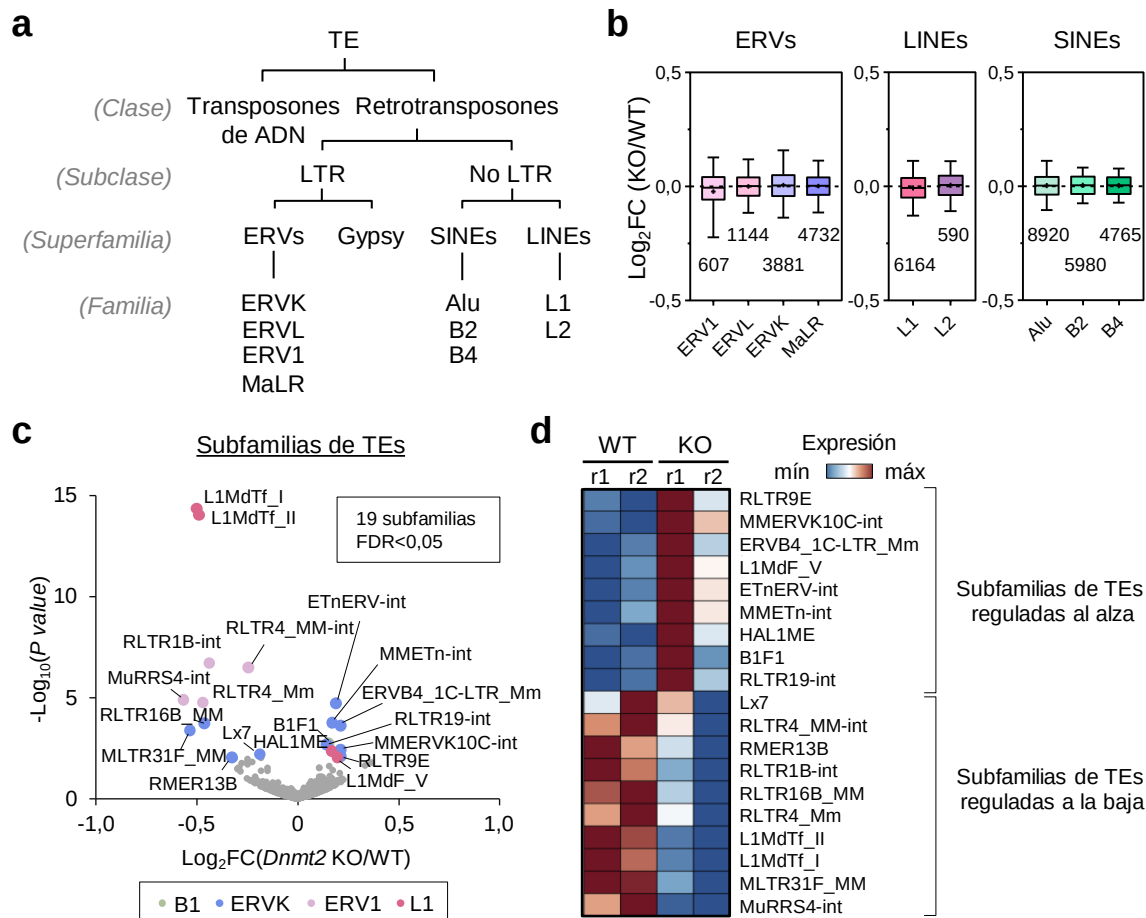


Figura R25. Análisis del efecto de la ausencia de DNMT2 en los niveles de expresión de TEs en ESCs. a) Clasificación de los elementos transponibles del genoma de mamíferos. b) Representación gráfica del cambio en la expresión de elementos de diversas familias de retrotransposones en ESCs *Dnmt2* KO en comparación con las WT. Los números indican los elementos de cada familia cuya expresión se evaluó. c) Representación por volcano plot del cambio en la expresión de diferentes subfamilias de retrotransposones en ausencia de DNMT2. Se destacan aquellas cuyo cambio en la expresión es estadísticamente significativo (FDR<0,05). El color indica la familia de retrotransposones a la que pertenecen (B1, ERVK, ERV1 o L1). d) Mapa de calor que ilustra la expresión de subfamilias de retrotransposones con cambios significativos (FDR<0,05) entre ESCs *Dnmt2* WT y KO.

6.2.3 La función de DNMT2 es necesaria para la salida de pluripotencia

Dado que nuestros datos sugieren que la pérdida de DNMT2 podría limitar el potencial de diferenciación de las ESCs, nos propusimos analizar las consecuencias de su ausencia durante la salida del estado celular pluripotente.

En primer lugar, llevamos a cabo un ensayo de generación de cuerpos embrioides (EBs, por sus siglas en inglés), en ESCs WT y KO para *Dnmt2*. Estas agrupaciones celulares tridimensionales permiten recapitular procesos tempranos del desarrollo embrionario en cultivo, favoreciendo la expresión de marcadores específicos de las tres capas germinales (ectodermo, mesodermo y endodermo) (Doetschman et al., 1985). Para llevar a cabo esta aproximación, mantuvimos las ESCs WT y KO para *Dnmt2* en cultivo en suspensión en un medio estándar pluripotente sin LIF durante un período de 10 días (**Figura R26a**).

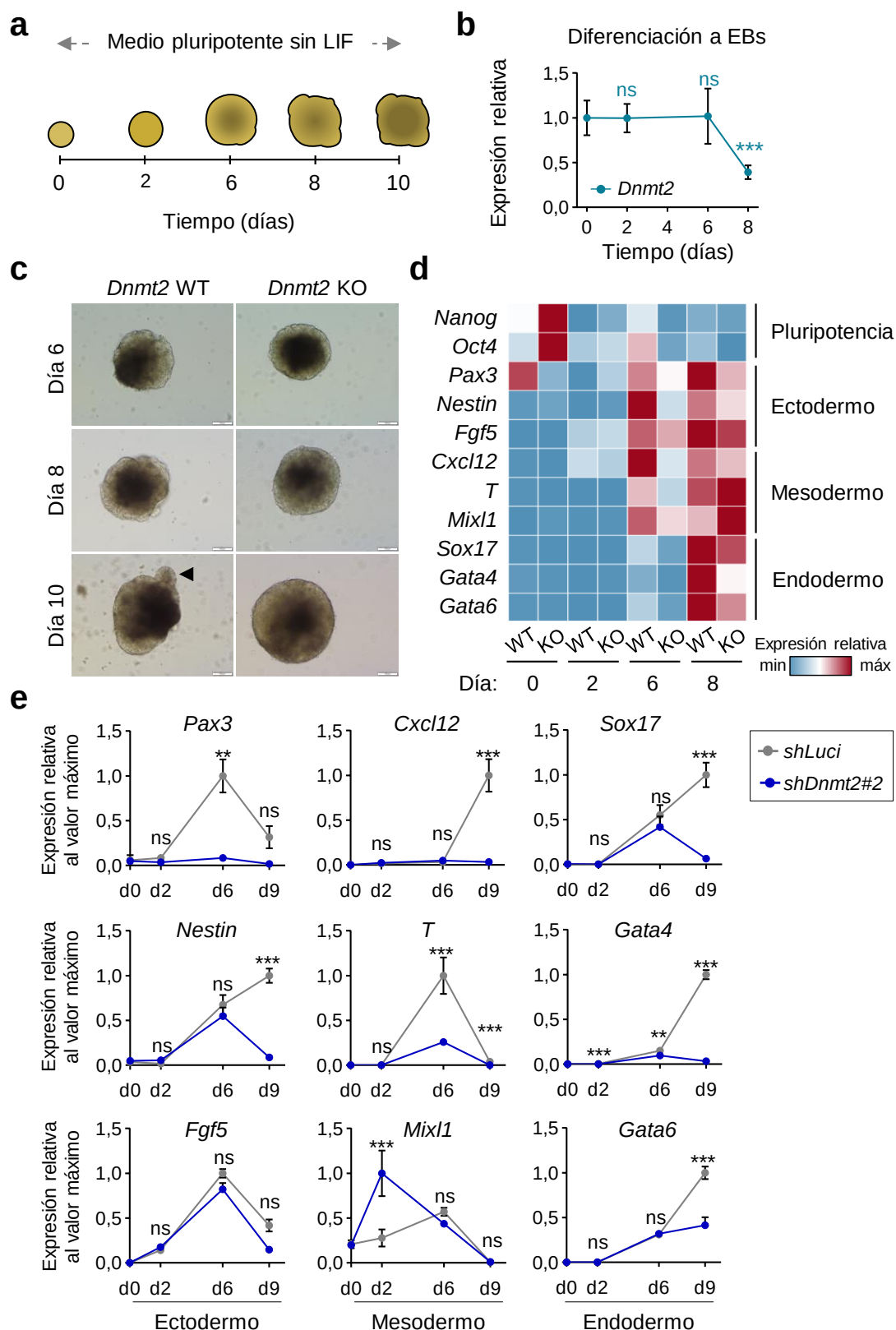


Figura R26. La presencia de DNMT2 es necesaria para la formación adecuada de EBs a partir de ESCs. a) Representación esquemática del proceso de diferenciación hacia EBs. Las ESCs WT y KO para *Dnmt2* se mantuvieron en cultivo durante 10 días en ausencia de LIF. **b)** Análisis de los niveles de expresión de *Dnmt2* por RT-qPCR durante la generación de EBs a partir de ESCs WT. **c)** Imágenes representativas de EBs generados a partir de ESCs WT y KO para *Dnmt2* en los días 6 (superior), 8 (central) y 10 (inferior). Se señala con flecha la evaginación

presente en los EBs WT a día 10 y ausente en los agregados celulares generados a partir de ESCs KO para *Dnmt2*. Escala: 100 μ m. **d)** Representación mediante mapa de calor de la expresión relativa de marcadores de pluripotencia y de las tres capas germinales (ectodermo, mesodermo y endodermo), determinada por RT-qPCR. **e)** Análisis mediante RT-qPCR de los niveles de expresión de marcadores de cada una de las tres capas germinales durante el proceso de formación de EBs a partir de ESCs transducidas con *shDnmt2#2* o *shLuci* (control). En b, d y e, la expresión se muestra con relación a *Hprt* y al día 0 (b y d) o a valor máximo de cada gen analizado (e). En b y e, los datos se representan como media $\pm$ desviación estándar de triplicados biológicos. Para el cálculo de diferencias estadísticamente significativas se aplicó una prueba ANOVA de un factor no pareada con corrección de Dunnet para comparaciones múltiples (b) y una prueba t de Student para cada día, con corrección FDR para comparaciones múltiples (e), indicadas como ns: no significativo, *: $p < 0,05$, **: $p < 0,01$ y ***: $p < 0,001$.

Cabe destacar que, en la generación de EBs bajo condiciones normales, los niveles de expresión de *Dnmt2* permanecen estables hasta disminuir después de 6 días en ausencia de LIF (**Figura R26b**). Esto sugiere el requerimiento de niveles relativamente elevados de *Dnmt2* en las primeras etapas del proceso de salida de pluripotencia. En consonancia con esto, observamos que los EBs generados a partir de ESCs KO para *Dnmt2* no mostraban las características evaginaciones presentes en los EBs formados a partir de ESCs WT después de 10 días en condiciones de diferenciación (**Figura R26c**). Además, el análisis mediante RT-qPCR de la expresión de marcadores de pluripotencia y específicos de las tres capas germinales confirmó una desregulación a nivel transcripcional en los EBs KO para *Dnmt2* en comparación con los WT (**Figura R26d**). Nuestros resultados indican que la presencia DNMT2 es necesaria para el adecuado silenciamiento de los marcadores de pluripotencia *Oct4* y *Nanog*, así como para la activación correcta de los marcadores de las tres capas germinales. Específicamente, se detectaron diferencias notables en la expresión de los marcadores de ectodermo *Pax3* y *Nestin*, los marcadores de mesodermo *Cxcl12* y *T* y los marcadores de endodermo *Gata4* y *Gata6* (**Figura R26d**). Asimismo, un análisis de expresión en EBs generados a partir de ESCs transducidas con shRNAs contra *Dnmt2* o *Luciferasa* respaldó estos resultados, ya que se detectó una regulación significativa a la baja de los marcadores *Pax3* y *Nestin* (ectodermo), *Cxcl12* y *T* (mesodermo) y *Sox17*, *Gata4* y *Gata6* (endodermo) (**Figura R26e**). Por lo tanto, estos datos resaltan la importancia de DNMT2 en la diferenciación apropiada de ESCs para establecer programas de transcripción específicos de las tres capas germinales.

Posteriormente, con el fin de confirmar los efectos de la pérdida de DNMT2 en la especificación de las capas germinales embrionarias durante la diferenciación temprana de ESCs, optamos por llevar a cabo un ensayo de generación de gastruloides en cultivo (Beccari et al., 2018). Los gastruloides son estructuras celulares tridimensionales que experimentan procesos similares a la gastrulación y elongación en un entorno de cultivo, replicando espacial y temporalmente la expresión génica temprana que tiene lugar después de la implantación del blastocisto en el útero (Beccari et al., 2018). Para su generación, ESCs WT y KO para *Dnmt2* fueron mantenidas en suspensión en pocillos con fondo en U en medio N2B27 durante 120 horas. Entre las 48 y 72 horas, añadimos al medio un inhibidor de la quinasa GSK3 β (CHIR99021), lo cual promueve la formación de agregados celulares, la elongación y la polarización de las estructuras (**Figura R27a**).

Los resultados revelaron acusados defectos morfológicos en los gastruloides generados a partir de ESCs *Dnmt2* KO en comparación con los WT (**Figuras R27b y R27c**). Específicamente, se observó una reducción significativa tanto en la eficiencia de generación de estas estructuras como en su tamaño en ausencia de DNMT2 (**Figuras R27d y R27e**). De modo similar al ensayo de generación de EBs, se analizaron los niveles de expresión de diversos marcadores, incluyendo aquellos de pluripotencia (*Oct4* y *Nanog*), diferenciación temprana (*Pax3*, *Cxcl12* y *Gata6*) e identidad posicional (*Hoxd3*, *Hoxd4* y *Hoxd13*) (**Figura R27f**). En

este caso, también encontramos una desregulación transcripcional en los genes de diferenciación temprana examinados, así como en los marcadores de identidad posicional *Hoxd3* y *Hoxd4* (**Figura R27f**).

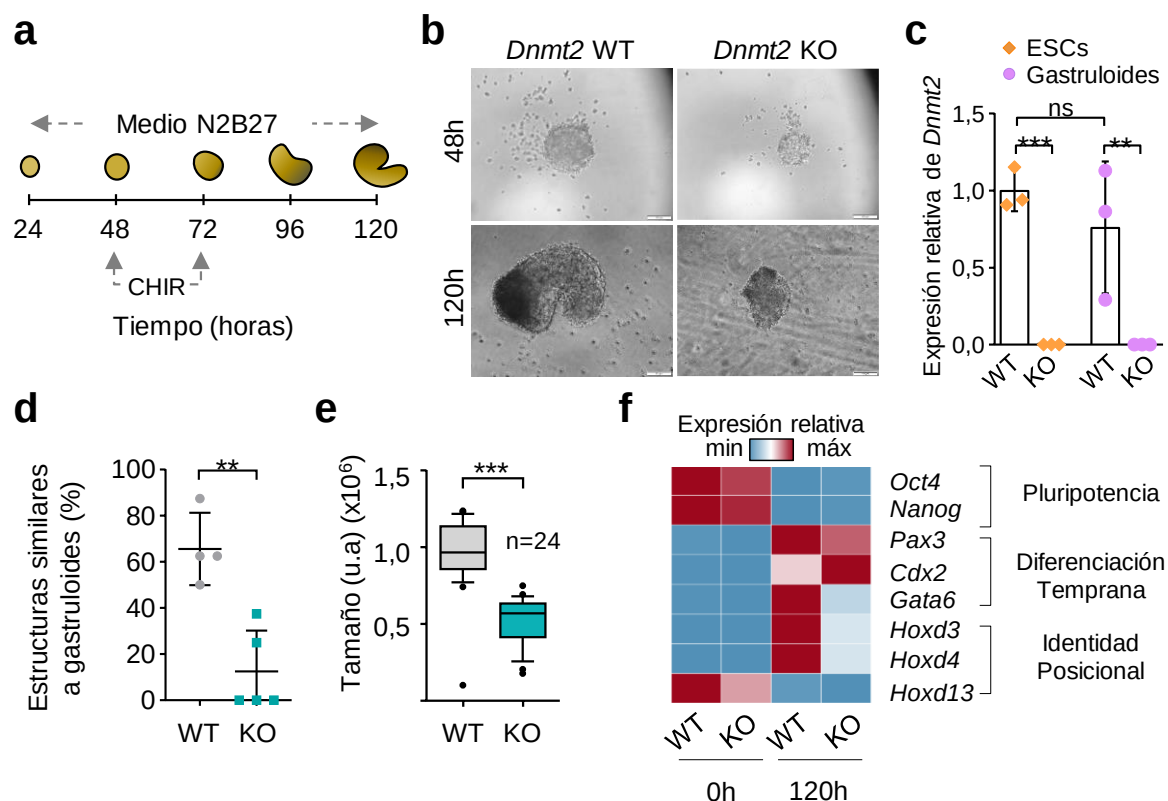


Figura R27. La ausencia de DNMT2 compromete la generación de gastruloides en cultivo. a) Esquema del proceso de generación de gastruloides. Se mantuvieron ESCs WT y KO para *Dnmt2* en cultivo en medio sin suero N2B27 en suspensión durante 120 horas, tratándose con CHIR entre las 48 y 72 horas del proceso. b) Imágenes representativas de las estructuras celulares tridimensionales generadas a partir de ESCs WT y KO para *Dnmt2* después de 48 y 120 horas en cultivo. Se observa la elongación, curvatura y polarización de las estructuras WT después de 120 horas, ausente en las estructuras *Dnmt2* KO. Escala: 100 μ m. c) Análisis por RT-qPCR de los niveles de expresión de *Dnmt2* en ESCs y gastruloides WT y KO para *Dnmt2*, relativos a *Hbms* y a ESCs WT. Los datos se representan como media $\pm$ desviación estándar de triplicados biológicos. d) Representación del porcentaje de estructuras similares a gastruloides generadas a partir de ESCs WT y KO para *Dnmt2*. Los datos se presentan como media $\pm$ desviación estándar de cuatro experimentos independientes, con seis agregados celulares tridimensionales analizados por experimento después de 120 horas en cultivo. e) Distribución del tamaño de las estructuras similares a gastruloides generadas en d. Las líneas en los extremos de las cajas se extienden al rango del percentil 10 al 90. La letra 'n' indica el número de agregados tridimensionales analizados. f) Mapa de calor que muestra la expresión de marcadores de pluripotencia, diferenciación temprana e identidad posicional en ESCs (0h) y agregados celulares tridimensionales (120h) *Dnmt2* KO en comparación con WT, determinada por RT-qPCR. La expresión se muestra en relación con *Hbms* y el valor máximo de cada gen analizado. En c, d y e, el cálculo de las diferencias estadísticas se llevó a cabo mediante una prueba ANOVA de un factor no pareada, con corrección de Sidak para comparaciones múltiples (c) y prueba t de Student de una cola no pareada (d y e), indicadas como ns: no significativo, **: $p < 0,01$ y ***: $p < 0,001$.

Colectivamente, estos resultados demuestran que la función de DNMT2 es necesaria para una formación eficiente de EBs y gastruloides en cultivo y confirman que su ausencia afecta la correcta inducción de marcadores de ectodermo, mesodermo y endodermo durante la salida de la pluripotencia.

nos propusimos investigar si la diferenciación de ESCs, que exhiben un fenotipo epitelial, podría verse afectada por la ausencia de DNMT2 en el proceso inverso, la transición epitelial-mesénquima (EMT), durante la diferenciación.

Con el fin de diferenciar ESCs WT y KO para *Dnmt2* a células comprometidas hacia el linaje mesénquima, estas se mantuvieron durante 5 días en condiciones estándar de cultivo. Durante este período, aplicamos factor de crecimiento transformante beta (TGF- β , por sus siglas en inglés) al medio de cultivo cada 24 horas, con o sin la presencia de LIF (**Figura R28a**). A nivel morfológico, se apreciaron diferencias entre las células WT y KO después de 5 días con TGF- β (**Figura R28b**). Además, notamos una disminución significativa en los niveles de la proteína marcadora α -SMA, que es característica de las células mesénquimas (**Figura R28c**). Es importante destacar que la presencia o ausencia de LIF no pareció influir en el proceso de diferenciación (**Figura R28c**).

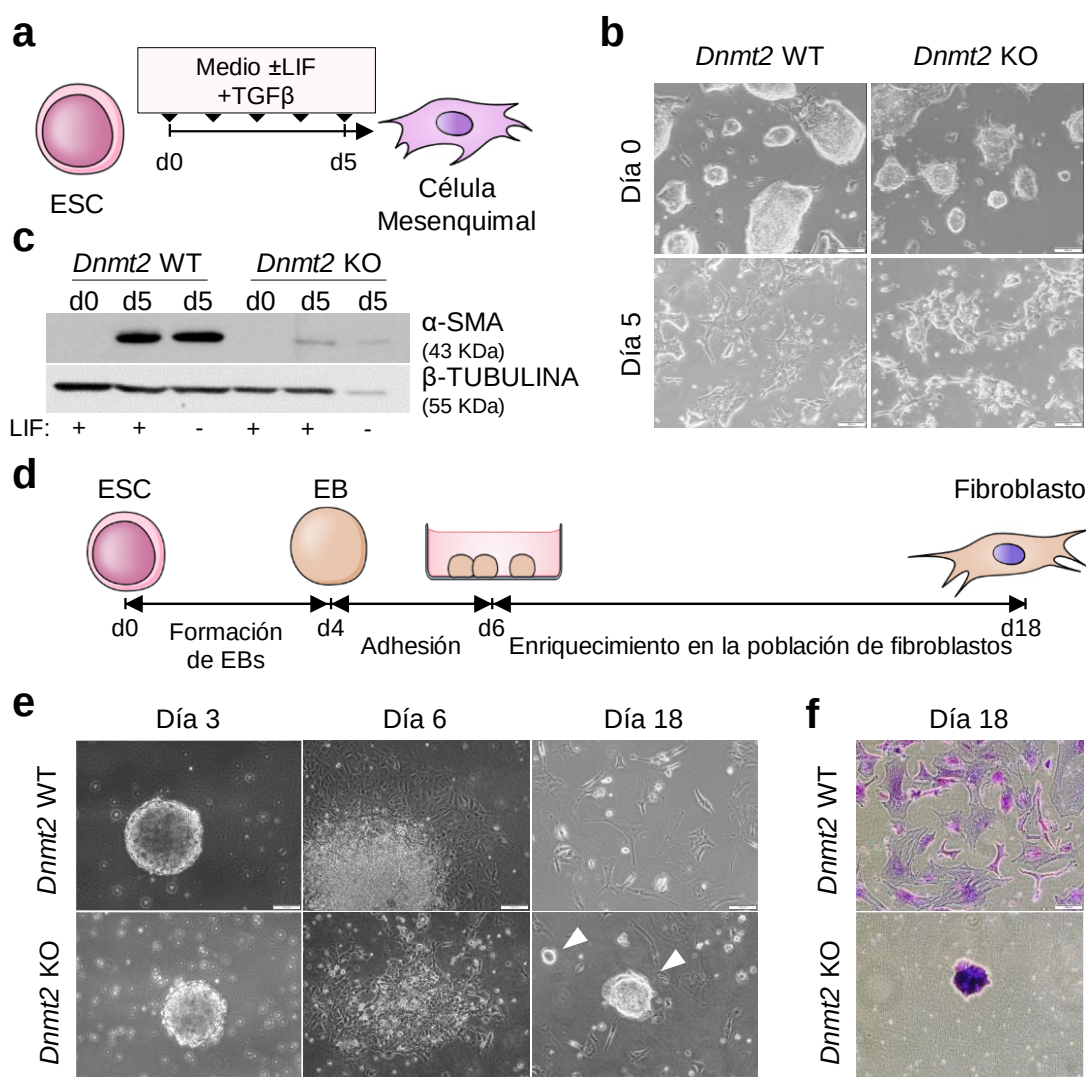


Figura R28. La falta de DNMT2 dificulta la diferenciación hacia el linaje mesénquima. **a)** Esquema de la estrategia de diferenciación de ESCs a células de tipo mesénquima. Las células se cultivaron durante cinco días en presencia de TGF- β [5 ng/ml]. **b)** Imágenes representativas de ESCs WT y KO para *Dnmt2* en los días 0 y 5 del proceso de diferenciación. **c)** Análisis mediante WB de los niveles de la proteína α -SMA en los días 0 y 5 del proceso de diferenciación en células WT y KO para *Dnmt2*. Se utilizan los niveles de β -TUBULINA como control de carga. **d)** Representación esquemática de la estrategia de diferenciación de ESCs a células tipo fibroblasto con un paso intermedio de generación de EBs. Las células se mantuvieron en suspensión durante 4 días en medio

sin LIF para favorecer la formación de EBs, que luego se adhirieron a placas de cultivo en presencia de gelatina al 0,1%. La disgregación posterior de los EBs facilitó el enriquecimiento en células de tipo fibroblasto. **e)** Imágenes representativas de células WT y KO para *Dnmt2* en los días 3, 6 y 18 del proceso de diferenciación a fibroblastos. Se señalan con flechas blancas agrupaciones celulares no diferenciadas tras 18 días en condiciones de diferenciación. **f)** Imágenes representativas después de la tinción con cristal violeta de células WT y KO para *Dnmt2* en el día 18 del proceso de diferenciación a fibroblastos. En b, d y f, la escala representa 100 μ m.

De manera paralela, llevamos a cabo un ensayo de diferenciación de ESCs a fibroblastos siguiendo un protocolo previamente publicado (Ruiz et al., 2012). Resumidamente, mantuvimos las ESCs en suspensión en medio sin LIF durante 4 días, lo que resultó en la formación de EBs. Posteriormente, los plaqueamos en pocillos tratados con gelatina al 0,1% para promover su adhesión. Después de dos días, procedimos a la individualización de las células, lo que llevó a un enriquecimiento en poblaciones celulares similares a fibroblastos, que se mantuvieron en cultivo durante 18 días (**Figura R28d**). El análisis morfológico reveló la incapacidad de las ESCs *Dnmt2* KO para alcanzar un estado completamente diferenciado, con la presencia de esferoides al cabo de 18 días de diferenciación (**Figura R28e**). Además, la tinción con cristal violeta reveló la presencia de una población homogénea de células similares a fibroblastos en la condición WT, pero no en el caso de las células *Dnmt2* KO (**Figura R28f**). Estos datos preliminares sugieren que DNMT2 podría desempeñar un papel importante en la toma de decisiones celulares requeridas para la EMT a partir de ESCs.

En resumen, los datos obtenidos resaltan la importancia de la función de DNMT2 en la toma adecuada de decisiones celulares que ocurren durante el proceso de diferenciación de ESCs.

6.2.4 La función m⁵C-metiltransferasa de DNMT2 en el ARN es necesaria en pluripotencia

Dado que DNMT2 es una m⁵C-metiltransferasa de ARN, decidimos explorar si sus funciones específicas en pluripotencia dependen o no de su actividad catalítica.

En primer lugar, analizamos los niveles de m⁵C en el ARN total de ESCs WT y KO para *Dnmt2* mediante un ensayo de inmunoabsorción ligada a enzima (ELISA, por sus siglas en inglés) (**Figura R29a**). Los resultados revelaron la ausencia de diferencias significativas en los niveles de m⁵C en el ARN total en ausencia de DNMT2 (**Figura R29b**).

Por otro lado, dado que el ARN ribosómico (ARNr) constituye la mayor parte de la fracción transcriptómica celular (Schaefer et al., 2017), con el fin de analizar posibles cambios en los niveles de m⁵C en el ARN mensajero (ARNm) y en el ARN largo no codificante (ARNlnc) después de la eliminación de DNMT2 en ESCs, procedimos a examinar los niveles de esta marca en la fracción larga de ARN tras la eliminación del ARNr. Este análisis se llevó a cabo mediante cromatografía líquida de alta resolución acoplada a espectrometría de masas (HPLC-MS/MS), para lo que el ARN total de ESCs *Dnmt2* WT y KO fue enriquecido en la fracción larga (>200 nucleótidos) y posteriormente tratado con el fin de eliminar el ARNr (**Figura R29c**). En este caso, tampoco detectamos diferencias significativas en los niveles de m⁵C después de la eliminación de DNMT2 (**Figura R29d**). Los resultados obtenidos concuerdan con un estudio reciente realizado por el equipo de la Dra. Michaela Frye en células HEK293T (Selmi et al., 2021), en el cual solo se observó una reducción significativa en los niveles de m⁵C en la fracción larga de ARN o en la fracción de ARN de transferencia (ARNt) después de silenciar simultáneamente DNMT2 junto con otras dos m⁵C-metiltransferasas de ARN: NSUN2 y NSUN6.

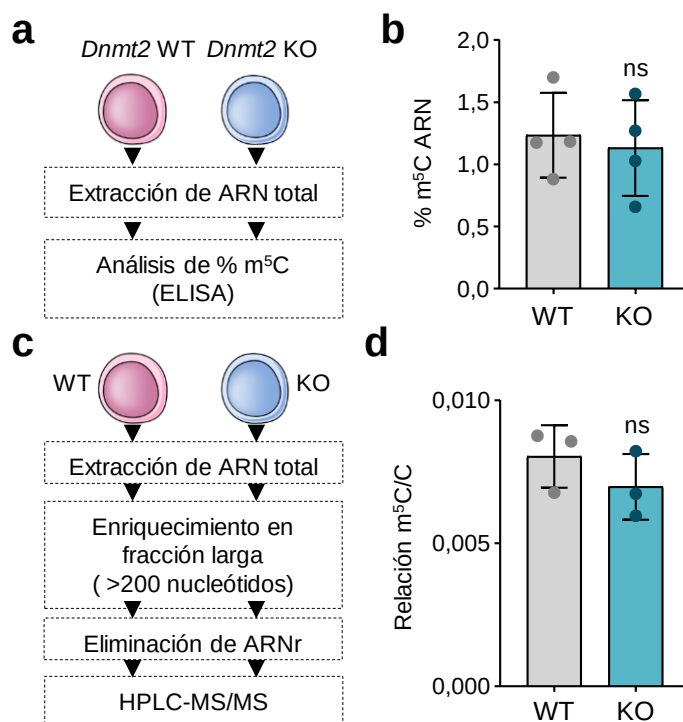


Figura R29. La ausencia de DNMT2 en ESCs no afecta a los niveles globales de m⁵C en ARN. a) Esquema del protocolo empleado para analizar los niveles globales de m⁵C en ARN total de ESCs *Dnmt2* WT o KO mediante ELISA. b) Cuantificación del porcentaje de m⁵C en ARN total evaluado por ELISA. c) Esquema del protocolo seguido para analizar los niveles globales de m⁵C en la fracción larga de ARN, después de la eliminación del ARNr, por HPLC-MS/MS en ESCs WT o KO para *Dnmt2*. d) Cuantificación de los niveles de m⁵C en la fracción larga de ARN, tras la eliminación del ARNr, detectados por HPLC-MS/MS. En b y d, los datos se presentan como media $\pm$ desviación estándar y el número de replicados biológicos se indica mediante puntos. Para el cálculo de las diferencias estadísticas se realizó una prueba t de Student de dos colas no pareada, indicadas como ns: no significativo.

Posteriormente, para indagar en el papel de DNMT2 en la distribución y los cambios de m⁵C a nivel de nucleótido en el transcriptoma durante la pluripotencia, empleamos conjuntos de datos previamente publicados de secuenciación de ARN tratado con bisulfito (RNA BS-seq). Esta metodología se basa en la desaminación de citosinas no metiladas mediante el tratamiento con bisulfito de sodio, convirtiéndolas en uracilos, mientras que las citosinas metiladas permanecen inalteradas. Posteriormente, la detección de citosinas metiladas puede ser determinada mediante diversas aproximaciones de secuenciación al comparar muestras tratadas y no tratadas con bisulfito de sodio. En este caso, seleccionamos y reanalizamos los datos de RNA BS-seq procedentes de dos estudios que documentan la presencia de la modificación m⁵C en los ARNs de ESCs de ratón (Amort et al., 2017; Legrand et al., 2017). Por un lado, en el estudio de Amort y colaboradores (Amort et al., 2017), se enfocaron en caracterizar los niveles de m⁵C en la fracción larga de ARN, para lo que realizan un enriquecimiento en aquellos transcritos que presentan cola de poli(A) (**Figura R30a**). Por otro lado, Legrand y colaboradores (Legrand et al., 2017), determinaron los niveles de m⁵C en tres fracciones de ARN, tanto en ESCs WT como KO para *Dnmt2*: i) fracción de ARNs pequeños (<200 nucleótidos), que principalmente contiene ARNr después de la eliminación del ARNr; ii) fracción larga de ARN, principalmente compuesta por ARNm y ARNlnc, tras la eliminación del ARNr; y iii) ARNr (**Figura R30a**). Teniendo esto en cuenta, y con el objetivo de establecer una lista fidedigna de moléculas de ARN modificadas con m⁵C en ESCs, procedimos a comparar los transcritos largos con m⁵C documentados en cada uno de los estudios seleccionados. Para ello, se procedió a la homogenización de los criterios de descubrimiento de

sitios metilados, mediante el análisis del conjunto de datos publicado por Legrand y colaboradores, aplicando los mismos criterios establecidos por Amort y colaboradores. Específicamente, seleccionamos los sitios metilados presentes en las tres réplicas biológicas y con un valor de FDR < 0,01 (**Figura R30b**).

En el estudio de Amort y colaboradores se identificaron 1.825 transcritos con colas de poli(A) que presentaban la modificación m⁵C en ESCs (Amort et al., 2017). Sin embargo, al volver a analizar los datos publicados por Legrand y colaboradores, de manera similar a lo que mostraron los autores, identificamos 270 transcritos largos modificados con m⁵C en ESCs *Dnmt2* WT (Legrand et al., 2017) (**Figura R30b**). Nuestro análisis comparativo reveló que solo 147 transcritos largos eran comunes a los dos estudios (**Figura R30b**).

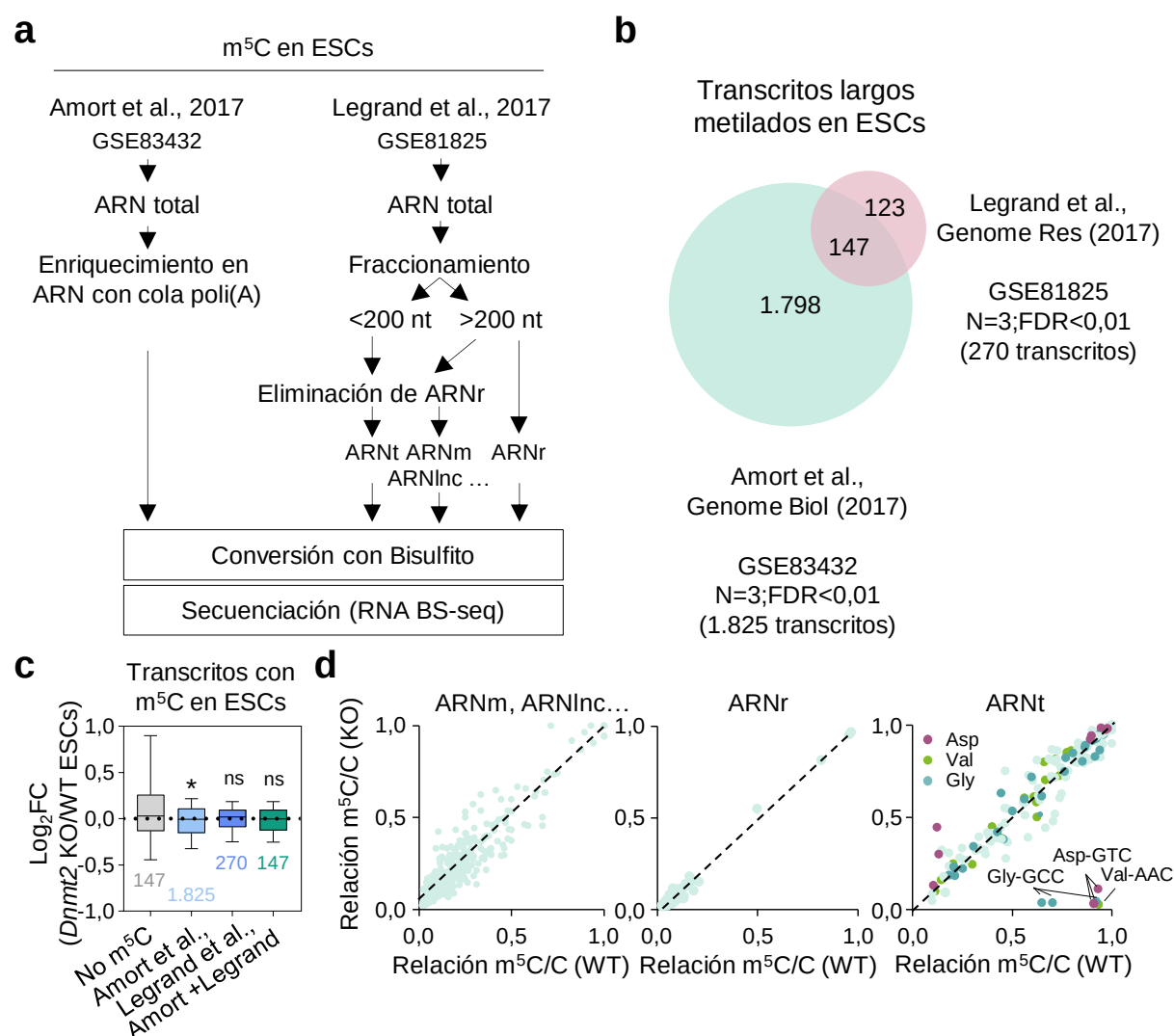


Figura R30. La falta de DNMT2 conduce a una pérdida específica de m⁵C38 en ARNt-Asp, ARNt-Gly y ARNt-Val. a) Esquema de las estrategias utilizadas por los dos estudios seleccionados para el análisis de m⁵C en ARN con resolución a nivel de nucleótido mediante RNA BS-seq en diversas fracciones de ARN en ESCs. Se indican los códigos GEO asociados a cada conjunto de datos tras su deposición. b) Comparación mediante diagrama de Venn de los transcritos largos con m⁵C en ESCs, detectados por los estudios seleccionados. Se indica con 'N' el número de replicados biológicos y el criterio de tasa de descubrimientos falsos empleado (FDR<0,01). c) Representación mediante diagrama de cajas de la expresión de los transcritos largos con m⁵C en ESCs basada en los datos obtenidos mediante RNA-seq de ESCs WT y KO para *Dnmt2*. Como control negativo, se seleccionaron de manera aleatoria 147 transcritos largos no metilados que se expresan en ESCs. Para la evaluación de diferencias significativas, se aplicó una prueba Kruskal-Wallis, indicada como ns: no significativo y *: p<0,05. d)

Representación mediante diagrama de puntos de los cambios de m⁵C en los tipos de ARN indicados en ESCs WT y KO para *Dnmt2*.

Por otro lado, con el fin de explorar si DNMT2 ejerce su función en ESCs a través de la metilación de transcritos largos, analizamos posibles alteraciones en la expresión de los transcritos con m⁵C en ESCs después de la pérdida de DNMT2, utilizando nuestros datos de RNA-seq de ESCs WT y KO para *Dnmt2*. Evaluamos los transcritos con m⁵C publicados en los estudios seleccionados (n=1.825 (Amort et al., 2017) y n= 270 (Legrand et al., 2017)), así como la lista de transcritos comunes a ambos (n=147). Observamos únicamente una reducción significativa en la expresión de los transcritos de la lista publicada por Amort y colaboradores tras la pérdida de DNMT2 (**Figura R30c**). Estos resultados sugieren que la eliminación de DNMT2 no afecta significativamente a los niveles de los transcritos largos con m⁵C en ESCs.

Luego, decidimos analizar el impacto de la ausencia de DNMT2 en los niveles de m⁵C en otros tipos de ARN. Para este propósito, aprovechamos los datos previamente publicados por Legrand y colaboradores (Legrand et al., 2017), de RNA BS-seq en ESCs WT y KO para *Dnmt2*. En consonancia con lo reportado previamente, no se observaron diferencias significativas en los niveles de m⁵C en los ARNs largos o en el ARNr entre las ESCs *Dnmt2* WT y KO (**Figura R30d**). Sin embargo, sí se observó una reducción significativa en los niveles de metilación en ciertos ARNt, específicamente en la citosina 38 (C38) de los ARNt-Gly-GCC, ARNt-Val-AAC y ARNt-Asp-GTC (**Figura R30d**). Esto concuerda con investigaciones previas que han identificado a los ARNt que transportan los aminoácidos valina (Val), glicina (Gly) y ácido aspártico (Asp) como las principales dianas de DNMT2 en su función de metilación del ARN (Goll et al., 2006; Jurkowski et al., 2008; Legrand et al., 2017; Schaefer et al., 2010; Tuorto et al., 2015; Tuorto et al., 2012; Zhang et al., 2018).

Finalmente, decidimos confirmar la reducción de los niveles de m⁵C en las moléculas de ARNt en ESCs tras la pérdida de DNMT2. Esta reducción había sido detectada previamente en nuestro análisis *in silico* de los datos de RNA BS-seq publicados por Legrand y colaboradores (Legrand et al., 2017), revelando una importante disminución en los niveles de m⁵C en la posición C38 (m⁵C38) de la molécula de ARNt-Asp-GTC, entre otras (**Figura R30d**). Este cambio en los niveles de m⁵C38 tras la eliminación de DNMT2 fue validado en el ARNt-Asp mediante conversión con bisulfito y posterior secuenciación Sanger (**Figuras R31a-R31c**). Así, nuestros resultados mostraron una pérdida prácticamente total de los niveles de m⁵C38 del ARNt-Asp en las ESCs KO para *Dnmt2* en comparación con las ESCs WT (**Figuras R31a-R31c**).

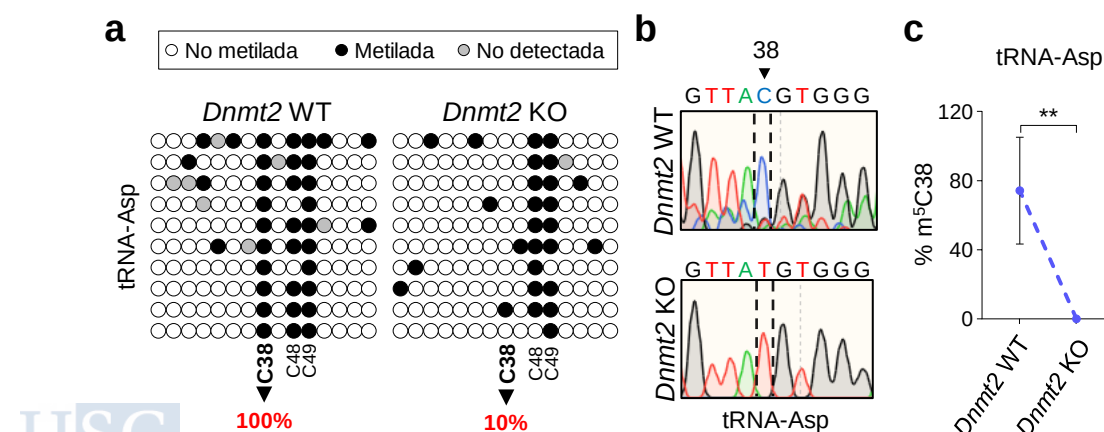


Figura R31. Validación de la actividad catalítica de DNMT2 en la posición C38 del ARNt-Asp en ESCs. a) Análisis RNA BS-seq de ARNt-Asp en ESCs WT y KO para *Dnmt2*. Después del tratamiento con bisulfito, el ARNt-Asp fue

amplificado por PCR y clonado en un vector para su secuenciación por el método de Sanger. Cada fila representa una réplica técnica ($n=10$). Se indican en blanco las citosinas no metiladas, en negro las metiladas y en gris aquellas posiciones en las que la citosina no fue detectada. Los números en rojo indican el porcentaje de metilación específica de la posición C38 (m^5C38), obtenido para cada condición. **b)** Secuencia del ARNt-Asp tras el tratamiento con bisulfito en ESCs WT o KO para *Dnmt2*. El triángulo invertido señala la posición C38. **c)** Cuantificación del porcentaje de metilación en la C38 obtenido para cada una de las condiciones analizadas. Los datos se muestran como media $\pm$ desviación estándar de triplicados biológicos. Para el cálculo de las diferencias estadísticamente significativas, se aplicó una prueba t de Student de una cola no pareada. Estas se indican como **: $p<0,01$.

En resumen, estos resultados respaldan que DNMT2 ejerce su actividad catalítica principalmente sobre las moléculas de ARNt y no sobre otros tipos de ARN en ESCs.

Posteriormente, con el fin de investigar la relevancia de la actividad catalítica de DNMT2 en pluripotencia y considerando la elevada homología (82%) entre las proteínas humana (hDNMT2) y de ratón (mDNMT2) (**Figuras R32a y R32b**), llevamos a cabo el rescate de la condición *Dnmt2* KO en ESCs mediante la expresión forzada de la versión humana. Las líneas de rescate generadas incluían la proteína humana ortóloga de DNMT2 (WT) y tres versiones mutantes (C79A, G155V y R371H) para la actividad m^5C en el ARN (Elhardt et al., 2015; Jurkowski et al., 2008) (**Figura R32c**). Cabe destacar que los vectores que contenían las secuencias codificantes para la proteína humana DNMT2 WT (hDNMT2) y las versiones mutantes (hDNMT2^{C79A}, hDNMT2^{G155V} y hDNMT2^{R371H}) fueron proporcionados por el laboratorio del Dr. Albert Jeltsch de la Universidad de Stuttgart. Estas secuencias fueron subclonadas en vectores de expresión del sistema PiggyBac de transposones, con tres epítomos Flag (3xFLAG) en su extremo N-terminal para facilitar su detección (**Figura R32c**).

Una vez generados los vectores de expresión, codificando tanto la proteína hDNMT2^{WT}, como las mutantes hDNMT2^{C79A}, hDNMT2^{G155V} y hDNMT2^{R371H}, se procedió a su transfección individual en la línea de ESCs *Dnmt2* KO, empleando un vector vacío (EV) como control (**Figura R32d**). Finalmente, procedimos a validar las líneas de ESCs generadas por WB (**Figura R32e**), lo que nos permitió detectar una reducción importante en los niveles de hDNMT2^{G155V} en comparación con la forma WT y las demás mutantes, lo cual ya había sido documentado previamente (Zhu et al., 2021). Esto sugiere una mayor inestabilidad de la proteína mutante hDNMT2^{G155V}, por lo que decidimos excluirla del análisis con el fin de no alterar las conclusiones extraídas en posteriores aproximaciones.

Específicamente, la proteína hDNMT2^{C79A} posee una mutación en el sitio catalítico correspondiente a un cambio de la cisteína en posición 79 (C79) por alanina (A), que impide la deposición de m^5C en el ARN (Jurkowski et al., 2008), mientras que la proteína hDNMT2^{R371H} presenta una actividad catalítica muy reducida (Elhardt et al., 2015). Con el fin de validar la reducción de la actividad catalítica de nuestras líneas de ESCs mutantes, se analizaron los niveles de metilación de la posición C38 del ARNt-Asp, mediante conversión del ARN con bisulfito y posterior secuenciación Sanger. El análisis confirmó la reducción de la actividad catalítica en las líneas de ESC KO mutantes, pero no en la línea KO rescatada con la forma WT de hDNMT2 (**Figura R32f**).

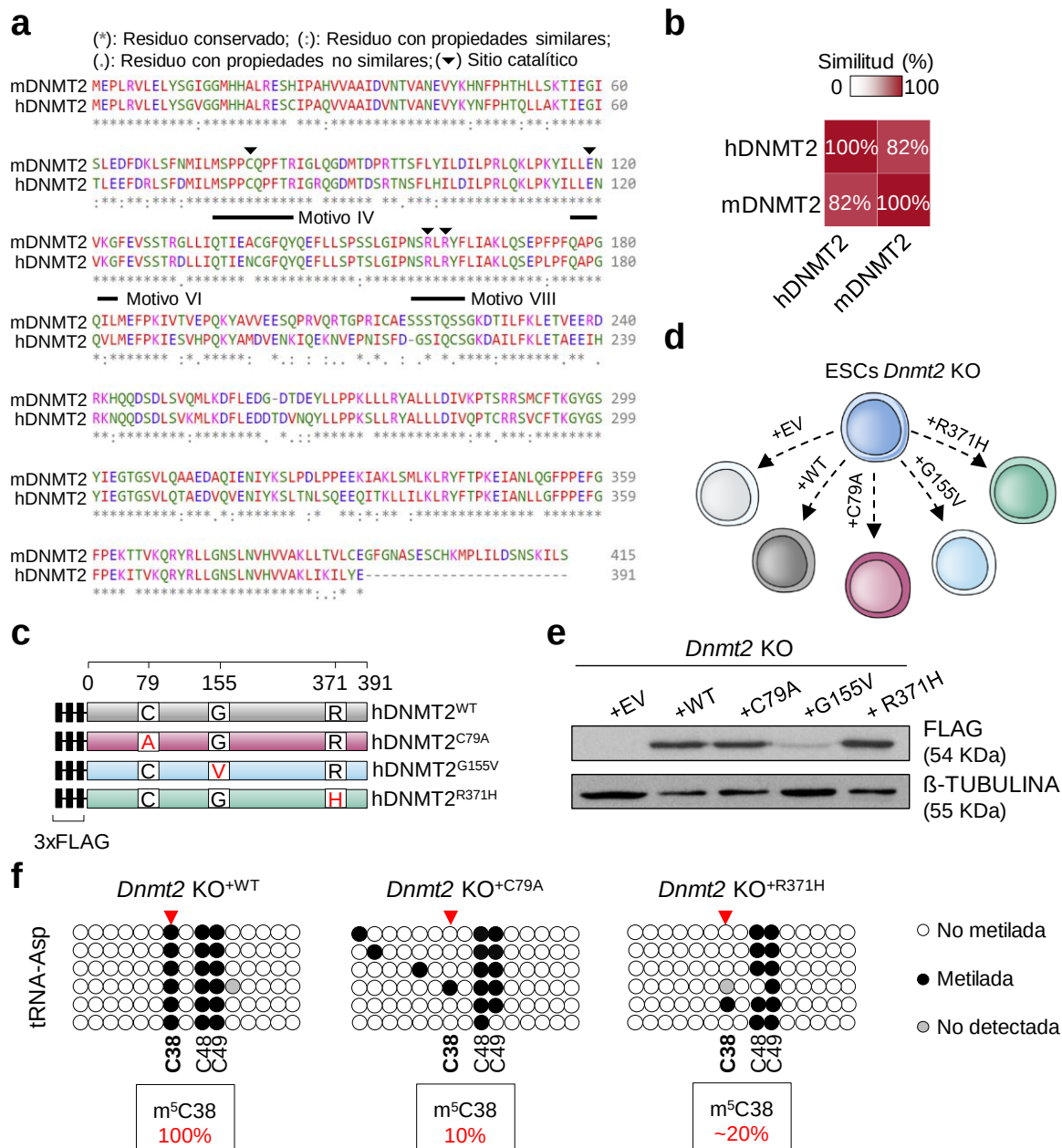


Figura R32. Generación y validación de líneas de rescate de DNMT2 WT y mutantes catalíticos en ESCs KO para *Dnmt2*. **a)** Alineamiento de las proteínas DNMT2 humana (hDNMT2) y de ratón (mDNMT2). Se indica el grado de conservación de los residuos como: residuo conservado (*), residuo no conservado con propiedades similares (:), y residuo no conservado con propiedades no similares. Se señalan los motivos proteicos (IV, VI y VIII) implicados en la actividad catalítica de DNMT2, altamente conservados entre ambas proteínas, así como los residuos que conforman el sitio activo (triángulo invertido). **b)** Matriz de similitud que indica el porcentaje de homología entre las secuencias proteicas de las versiones humana y murina de DNMT2. **c)** Esquema de la proteína DNMT2 humana WT y de las versiones mutantes catalíticas (C79A, G155V y R371H). Se indican los tres epítopos Flag (3xFLAG) añadidos al extremo N-terminal, los cambios de aminoácidos en las proteínas mutantes (rojo) y su posición. **d)** Esquema de las líneas de rescate de DNMT2 WT y mutantes catalíticos en las ESCs KO para *Dnmt2*. Se empleó la transfección con el vector vacío (EV) como control negativo. **e)** Análisis mediante WB de los niveles de las versiones WT y mutantes de hDNMT2 en las líneas estables de rescate generadas, detectados mediante un anticuerpo contra el epítipo FLAG. Se muestran los niveles de β-TUBULINA como control de carga. **f)** Detección por secuenciación Sanger de ARN tratado con bisulfito de m⁵C38 en el ARNt-Asp de ESCs *Dnmt2* KO rescatadas con DNMT2 WT o las versiones catalíticas mutantes C79A y R371H. Cada fila representa una réplica técnica (n=6). Se indican en blanco las citosinas no metiladas, en negro las metiladas y en gris aquellas posiciones en las que la citosina no fue detectada. Los números en rojo indican el porcentaje de metilación específico de la

posición C38 (m⁵C38), obtenido para cada condición. Se indican los niveles de m⁵C en las citosinas 48 y 49, dianas reportadas de NSUN2, como control.

Después, con el fin de evaluar el impacto de la pérdida de la actividad catalítica de DNMT2 en el transcriptoma de las células pluripotentes, realizamos un análisis mediante RNA-seq en las líneas de ESCs KO para *Dnmt2* rescatadas con las versiones mutantes hDNMT2^{C79A} y hDNMT2^{R371H}. En primer lugar, mediante un PCA pudimos confirmar que el perfil de expresión génica de las líneas mutantes catalíticas (*Dnmt2* KO^{+C79A} y *Dnmt2* KO^{+R371H}) guarda una mayor similitud con el de las ESCs *Dnmt2* KO que con el de las WT (**Figura R33a**). Además, aplicando el coeficiente de correlación de Pearson pudimos generar una matriz de similitud acoplada a una agrupación jerárquica, que nos permitió clasificar diferencialmente el transcriptoma de las ESCs WT, KO y las rescatadas con las versiones mutantes catalíticas inactivas (**Figura R33b**).

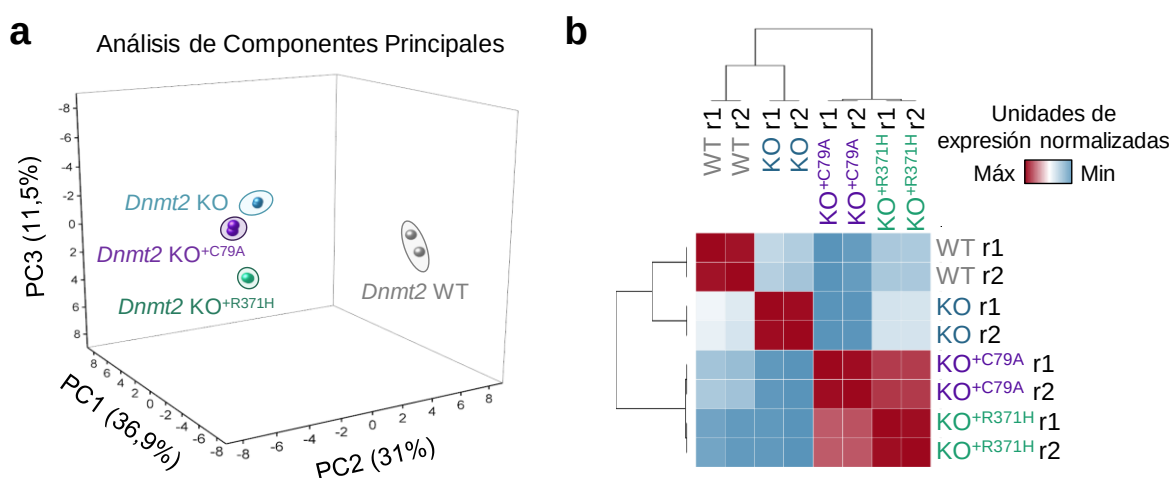


Figura R33. La pérdida de la actividad m⁵C-metiltransferasa de DNMT2 altera el perfil transcripcional de las ESCs. a) PCA de los perfiles transcripcionales de las líneas de ESCs indicadas. El porcentaje de variación explicado por cada PC se indica en el eje correspondiente. b) Análisis de los perfiles transcripcionales de las líneas de ESCs indicadas mediante una matriz de similitud acoplada a una agrupación jerárquica en base al coeficiente de correlación de Pearson.

Para determinar aquellos transcritos cuya expresión depende de la actividad catalítica de DNMT2, analizamos la huella transcripcional (FDR<0,05 y >1,5FC) generada en ESCs tras la eliminación de DNMT2 (**Figura R24a**), en las líneas mutantes catalíticas *Dnmt2* KO^{+C79A} y *Dnmt2* KO^{+R371H}. De este modo, pudimos clasificar los transcritos significativamente desregulados en ausencia de DNMT2 en dos grupos (**Figura R34a**). Por un lado, el grupo I está constituido por aquellos transcritos (n=214) cuya regulación es independiente de la actividad catalítica de DNMT2, puesto que la expresión forzada de las versiones mutantes catalíticas de la proteína es suficiente para rescatar su expresión a niveles similares a los de las ESCs WT (**Figura R34a**). Por otro lado, el grupo II se encuentra integrado por los transcritos (n=120) cuya regulación sí depende de la actividad catalítica de DNMT2, ya que la expresión forzada de proteínas DNMT2 mutantes sin actividad catalítica en ESCs *Dnmt2* KO no es suficiente para recuperar los niveles de expresión que presentan las ESCs WT (**Figura R34a**). Dentro cada uno de los grupos establecidos podemos distinguir dos clases de transcritos. El grupo I (independiente de actividad catalítica) estará conformado tanto por los 167 genes regulados a la baja en ausencia de DNMT2 (Clase I), como por los 47 genes regulados al alza en ausencia de DNMT2 (Clase II). Lo mismo ocurre en el caso del grupo II (dependiente de actividad catalítica), compuesto por 23 genes regulados al alza tras la pérdida de DNMT2 (Clase III) y 97 genes regulados a la baja (Clase IV) (**Figura R34a**).

A continuación, se procedió a determinar en qué procesos biológicos estaban involucrados estos grupos de transcritos. Para lograrlo, realizamos un análisis de enriquecimiento en ontologías génicas empleando la herramienta en línea PANTHER (**Figura R34b**). Los resultados revelaron que el grupo I de genes, cuya regulación transcripcional no depende de la actividad catalítica de DNMT2, se encuentra asociado con procesos de angiogénesis, apoptosis y migración celular, entre otros (**Figura R34b**). En cambio, el grupo II de genes, cuya expresión sí es dependiente de la actividad catalítica de DNMT2, se relaciona con procesos de expresión génica y diferenciación, particularmente relevantes en eventos de salida de pluripotencia (**Figura R34b**). Por lo tanto, estos resultados sugieren que la pérdida de actividad catalítica de DNMT2 en ARN podría estar involucrada en los defectos observados en las ESCs durante la salida de pluripotencia (**Figuras R26-R28**).

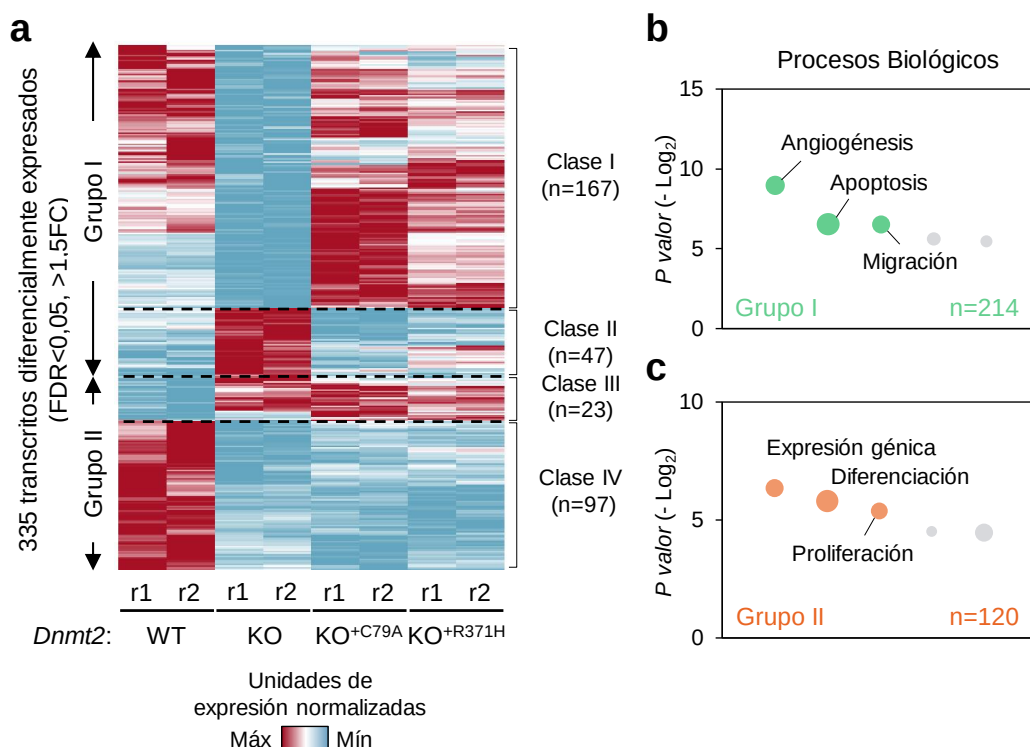


Figura R34. DNMT2 interviene en la regulación de programas transcripcionales en ESCs mediante funciones dependientes e independientes de su actividad catalítica. a) Evaluación de la huella transcripcional (FDR<0,05, FC>1,5) en ESCs tras la eliminación de DNMT2, en células WT, KO y rescatadas con las versiones mutantes catalíticas (C79A y R371H). Los genes significativamente desregulados se clasifican en dos grupos: aquellos cuya regulación no depende de la actividad catalítica de DNMT2 (grupo I) y los que sí dependen de esta actividad (grupo II). Estos grupos se subdividen según si los transcritos están regulados a la baja (clases I y IV) o al alza (clases II y III) en ausencia de DNMT2. Análisis de los procesos biológicos, mediante PANTHER, en los que participan los genes pertenecientes al grupo I (b) y al grupo II (c). Se indica el p valor (-Log₂) y el número de transcritos pertenecientes a cada grupo.

Para evaluar la importancia de la actividad catalítica de DNMT2 en la salida de pluripotencia, llevamos a cabo un ensayo de diferenciación celular de ESCs hacia los linajes de neuroectodermo (NE) y mesendodermo (ME) (Thomson et al., 2011). Utilizamos ESCs tanto WT como KO para *Dnmt2*, así como ESCs KO rescatadas con las versiones humanas WT y la mutante C79A. Durante el ensayo, las células se cultivaron en medio sin suero (N2B27), con la adición de LIF, el inhibidor de la quinasa MEK (PD0325901) y la citocina BMP4 en las primeras 48 horas. Posteriormente, se mantuvieron por 48 horas adicionales en medio N2B27 y, finalmente, se administraron durante 72 horas el inhibidor de la quinasa GSK3β

(CHIR99021) para la diferenciación a ME, o ácido retinoico (RA) para la diferenciación a NE (**Figura R35a**).

Previamente al inicio del ensayo, validamos por RT-qPCR la expresión del gen *DNMT2* humano (*hDNMT2*) en las ESCs *Dnmt2* KO sobre las que se había inducido la expresión de las proteínas humanas WT o mutante C79A (**Figura R35b**).

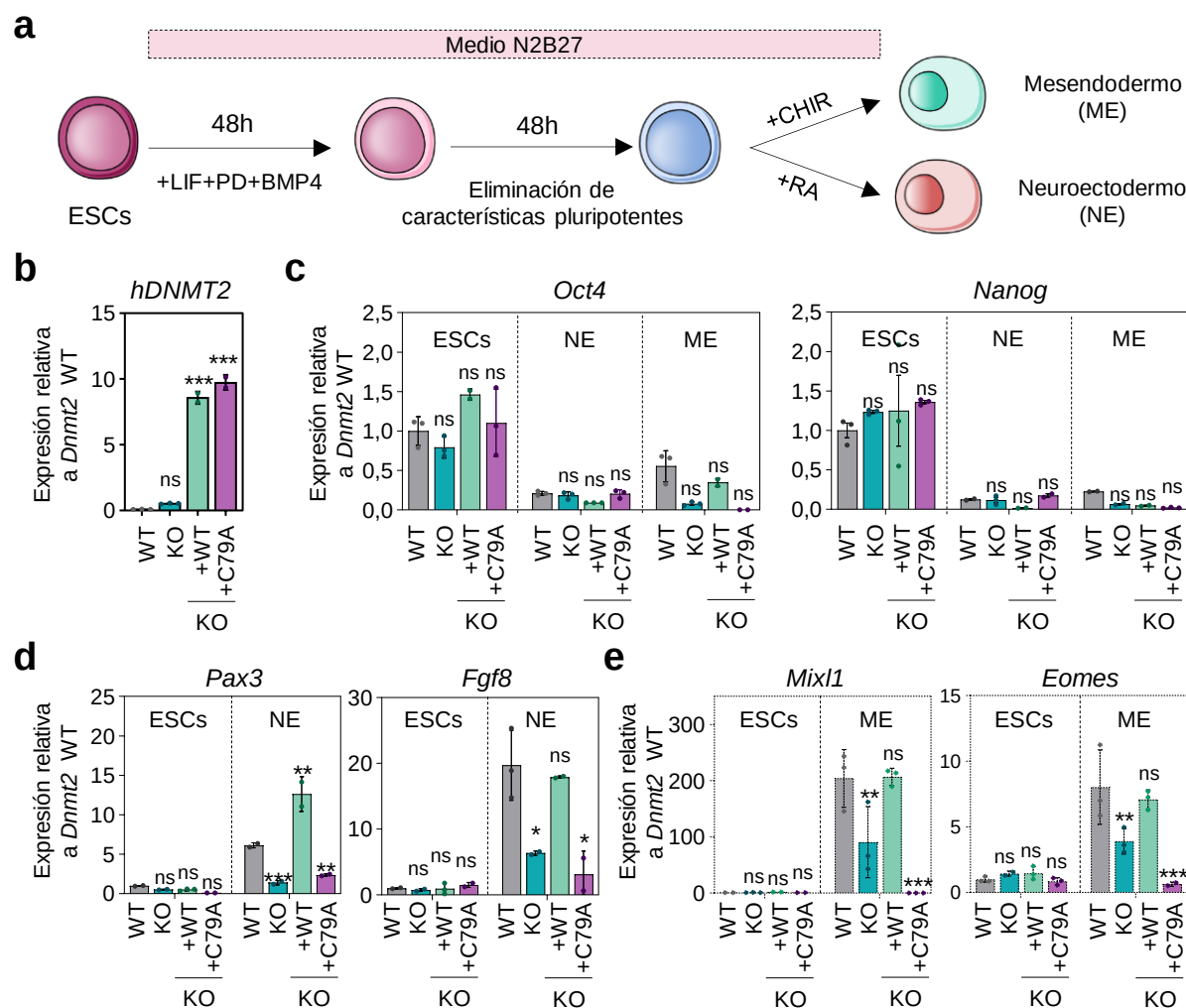


Figura R35. La actividad catalítica de DNMT2 es crítica en eventos de especificación temprana. a) Estrategia de diferenciación de ESCs hacia los linajes mesendodermo (ME) y neuroectodermo (NE). b) Análisis de los niveles de expresión del gen *hDNMT2* en ESCs WT o KO para *Dnmt2*, así como en ESCs KO que expresan las versiones humanas WT o mutante catalítica C79A. Niveles de los marcadores de pluripotencia (*Oct4* y *Nanog*) (c), marcadores de NE (*Pax3* y *Fgf8*) (d), y marcadores de ME (*Mixl1* y *Eomes*) (e), en las condiciones indicadas. En b, c, d y e, los datos fueron obtenidos mediante RT-qPCR y se presentan como media $\pm$ desviación estándar, normalizados respecto a *Hprt* y a ESCs WT. El número de replicados biológicos se indica mediante puntos. La significancia de las diferencias observadas fue evaluada mediante una prueba ANOVA de dos factores no pareada con corrección de Dunnett para comparaciones múltiples, indicadas como ns: no significativo, **: $p < 0,01$ y ***: $p < 0,001$.

Después de la diferenciación de las ESCs hacia los distintos linajes, se determinó que en todas las condiciones se redujo la expresión de los marcadores de pluripotencia *Oct4* y *Nanog* en comparación con las ESCs (**Figura R35c**). De acuerdo con los resultados previos (**Figuras R26d, R26e y R27f**), la ausencia de DNMT2 resultó en una reducción significativa de la expresión de los marcadores *Pax3* y *Fgf8* durante la diferenciación de las ESCs hacia el

linaje NE (**Figura R35d**). Este defecto únicamente fue revertido tras la expresión de la forma catalíticamente activa de DNMT2 ($hDNMT2^{WT}$) y no de la forma mutante catalítica ($hDNMT2^{C79A}$) (**Figura R35d**). De manera similar, durante la diferenciación hacia ME, solo la expresión de la proteína $hDNMT2^{WT}$ y no la de la mutante $hDNMT2^{C79A}$, pudo atenuar la reducción en la expresión de los marcadores *Mixl1* y *Eomes* observada tras la diferenciación de las ESCs *Dnmt2* KO (**Figura R35e**).

En conjunto, estos resultados apoyan que la actividad m^5C -metiltransferasa de DNMT2 participa en la correcta diferenciación celular de las ESCs.

6.2.5 Identificación de los interactores de DNMT2 en ESCs

Con el fin de explorar la función independiente de actividad catalítica de DNMT2 en la regulación del estado celular pluripotente, decidimos llevar a cabo la identificación de las proteínas con las que interactúa en ESCs, tanto de manera directa como indirecta. Para llevar a cabo esta tarea, empleamos una estrategia de inmunoprecipitación (IP) para la purificación por afinidad de los complejos proteicos unidos a DNMT2 en ESCs, en combinación con espectrometría de masas (**Figura R36a**).

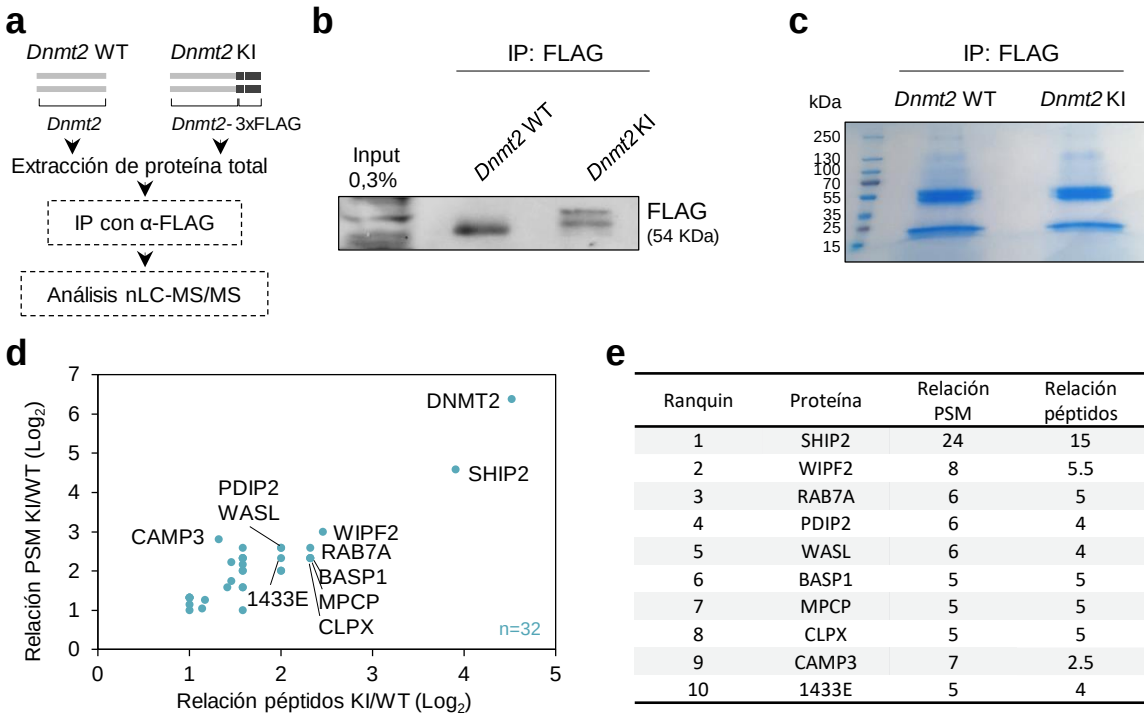


Figura R36. Identificación del interactoma de DNMT2 en ESCs. a) Representación esquemática de la estrategia utilizada para identificar las proteínas que interactúan con DNMT2 en ESCs. b) Análisis mediante IP y WB de los niveles de la proteína DNMT2-3xFLAG después de la IP con α -FLAG, detectados en la línea de ESCs *Dnmt2* KI y no en la *Dnmt2* WT. c) Electroforesis de las IPs indicadas en un gel Tris-Glicina de acrilamida/bisacrilamida con gradiente del 4% al 20%. d) Representación gráfica de los 31 interactores de DNMT2 identificados por nLC-MS/MS. Se señalan DNMT2 y los interactores con mayor puntuación según los criterios establecidos. e) Lista de las 10 proteínas con la mayor probabilidad de interaccionar con DNMT2 en ESCs, considerando las relaciones de péptidos y el número total de secuencias peptídicas (PSM) entre las ESCs *Dnmt2* KI y WT.

En primer lugar, se añadieron tres epítopos Flag (3xFLAG) al extremo C-terminal de la proteína DNMT2 editando cada uno de los alelos del gen que la codifica mediante la tecnología CRISPR/Cas9, lo que resultó en una línea *knockin* (KI) homocigótica: *Dnmt2* KI. La validación de esta línea se realizó mediante secuenciación Sanger, comprobándose posteriormente que sus

niveles proteicos eran comparables a los de las ESCs parentales (WT) mediante WB. Luego, extrajimos proteína total de ESCs WT, como control, y ESCs KI para después realizar una IP utilizando un anticuerpo α -FLAG (**Figura R36b**). Los complejos proteicos inmunoprecipitados en ambas condiciones se separaron por tamaño (**Figura R36c**) y se analizaron mediante espectrometría de masas en tándem por cromatografía de nanolíquidos (nLC-MS/MS) (**Figura R36d**).

Para garantizar una identificación lo más precisa posible de las proteínas que interactúan con DNMT2 en ESCs, aplicamos los siguientes criterios: i) tanto el número de péptidos como las secuencias peptídicas (PSM) identificadas en la muestra KI deben ser iguales o superiores a 1 y ii) tanto la relación del número de péptidos distintos como la relación de PSM de cada potencial interactivo de DNMT2 deben ser iguales o mayores a 2, en la línea KI con respecto a la WT. De este modo, identificamos 31 proteínas con una elevada confianza de ser interactores reales de DNMT2 (**Figuras R36d y R36e**), destacando especialmente aquellas que se describen en la **Figura R36e**, como las más probables.

Para evaluar el comportamiento de los genes que codifican las proteínas que interactúan con DNMT2, nos propusimos determinar si su expresión se encontraba alterada tras la pérdida de DNMT2. Con este fin, se integraron los datos del RNA-seq de ESCs WT y KO para *Dnmt2* y los datos del interactoma, utilizando el programa Cytoscape. Nuestro análisis reveló que la expresión de 15 de los genes que codifican para las 31 proteínas que interactúan con DNMT2 (equivalente al 48,4%) estaba significativamente alterada ($FDR < 0,05$) en las ESCs KO para *Dnmt2* (**Figura R37a**). Sin embargo, es importante destacar que la tasa de cambio en la expresión de estos genes fue inferior al 50% en cada uno de ellos (**Figura R37b**).

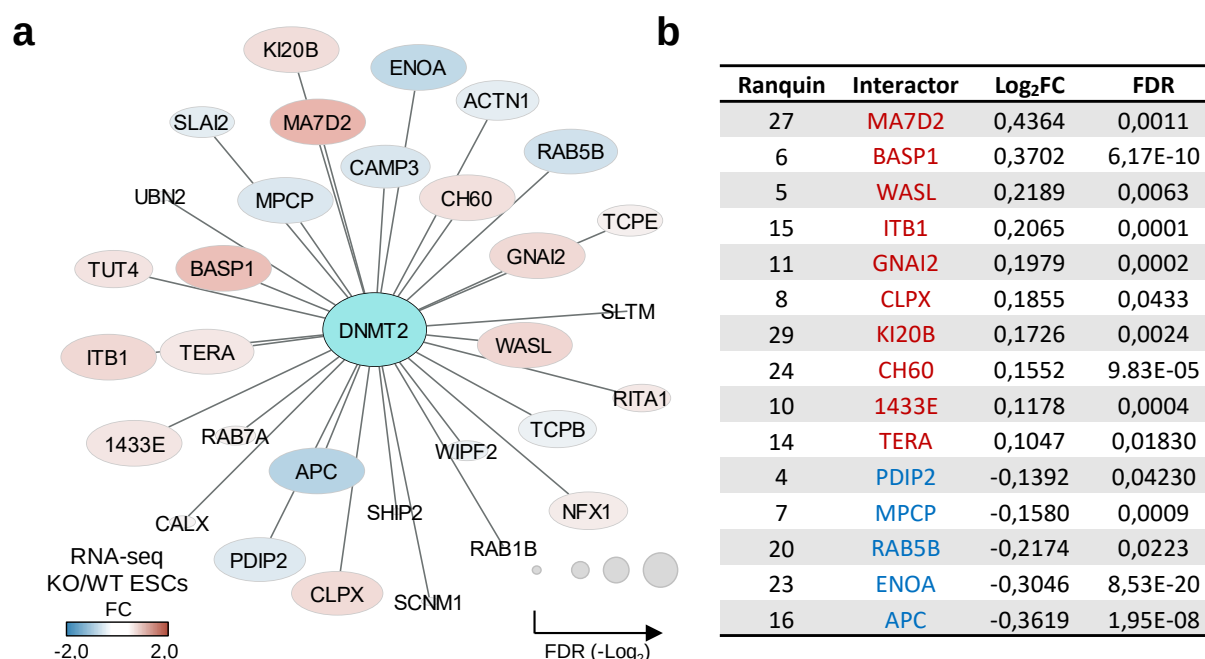


Figura R37. Evaluación del impacto de la falta de DNMT2 en la expresión de genes que codifican sus proteínas interactoras. a) Representación de las 31 proteínas que interactúan con DNMT2 en ESCs, identificadas por nLC-MS/MS. Se muestra el cambio en la expresión (FC) de los genes que las codifican tras la pérdida de DNMT2. El tamaño de la burbuja refleja el valor de significancia estadística (-Log₂ FDR). b) Tabla mostrando las proteínas cuya expresión en los genes que las codifican se ve significativamente ($FDR < 0,05$) incrementada (en rojo) o reducida (en azul) en ausencia de DNMT2. Se proporciona la posición que cada proteína ocupa en el ranquin de interactores de DNMT2, junto con el FC y el FDR relativos a su alteración transcripcional.

Posteriormente, nos propusimos explorar los procesos biológicos en los que intervienen los interactores de DNMT2. Para ello, realizamos una búsqueda exhaustiva en la literatura científica y en diversas bases de datos, como STRING, UniProt, GeneCards, entre otras. De este modo, pudimos identificar la implicación de las proteínas que interactúan con DNMT2 en una amplia diversidad de procesos biológicos que abarcan desde la organización celular y la transcripción y procesamiento del ARN, hasta la respuesta a daños en el ADN, el ciclo y la migración celulares, entre otros (**Figura R38a**).

Además, para evaluar las posibles consecuencias de la alteración de las proteínas interactoras de DNMT2, llevamos a cabo un análisis de ontología de fenotipos de mamífero (MPO), empleando la herramienta en línea Enrichr. Nuestros resultados indicaron que la disrupción de los interactores de DNMT2 está relacionada con un menor tamaño del embrión, una disfunción en la miogénesis y un aumento en la incidencia de tumores mamarios, entre otros (**Figura R38b**).

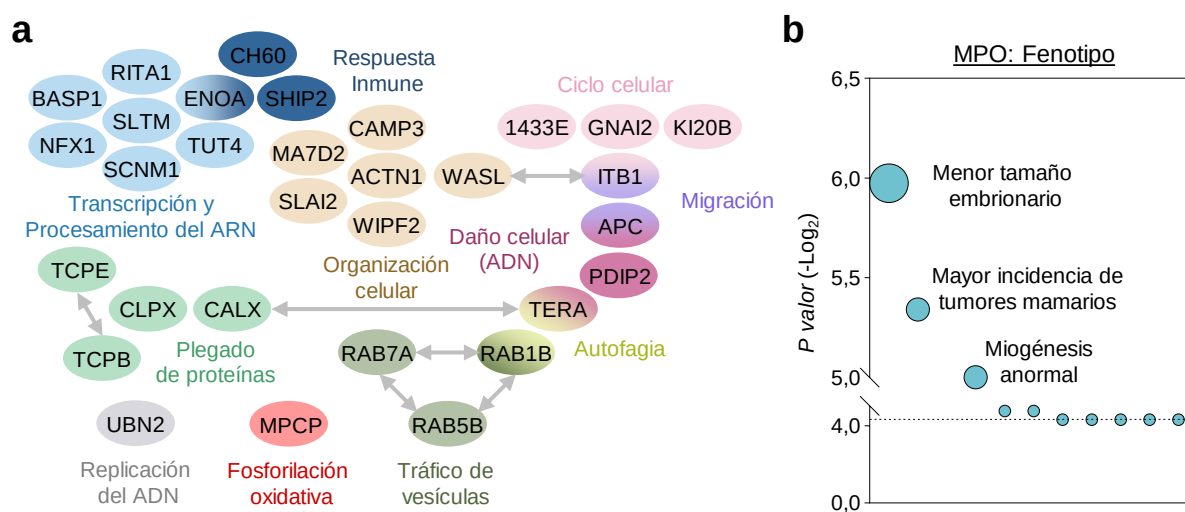


Figura R38. Análisis de los procesos y fenotipos asociados a los interactores de DNMT2. a) Diagrama que muestra los procesos biológicos en los que se encuentran implicados los interactores de DNMT2 en ESCs. Cada color indica la implicación de las proteínas en distintos procesos, mientras que la interacción entre proteínas es señalada mediante flechas. b) Análisis de ontología de fenotipos de mamíferos (MPO) empleando una lista que contiene a DNMT2 y sus 31 interactores. Se muestra el enriquecimiento de tres de los diez principales términos identificados, señalando mediante una línea de puntos el umbral de significancia FDR<0,05.

Estos hallazgos sugieren que la alteración de la función no dependiente de la actividad catalítica de DNMT2 podría tener implicaciones en el desarrollo, el crecimiento y eventos de transformación oncogénica (**Figuras R38a y R38b**). Sin embargo, no podemos descartar la posibilidad de que algunas de las proteínas con las que DNMT2 interactúa en ESCs sean necesarias para regular su actividad catalítica en ARN. Dada la relevancia de estos resultados, será esencial llevar a cabo futuros estudios para profundizar en esta nueva línea de investigación y comprender mejor los mecanismos subyacentes.

6.3 EXPLORACIÓN DEL PAPEL DE DNMT2 EN CÁNCER

6.3.1 Las mutaciones en el gen *DNMT2* o la reducción de sus niveles de expresión se asocian con un peor pronóstico en varios tipos de cáncer

Tal y como hemos descrito previamente (ver **Figura R23**), nuestros resultados en ESCs sugieren que DNMT2 podría regular los niveles de la proteína P53 a través de la regulación transcripcional de *Mdm2*.

Numerosas evidencias respaldan las consecuencias de mutaciones y cambios en los niveles del gen que codifica P53 en varios tipos de cáncer (Muller y Vousden, 2014; Sabapathy y Lane, 2018). Teniendo esto en cuenta, y que DNMT2 se relaciona en ESCs con procesos biológicos comúnmente alterados en eventos de transformación oncogénica, como la angiogénesis, la migración, la apoptosis y la proliferación, entre otros (ver **Figura R34b**), además de con un mayor riesgo de desarrollar ciertos tipos de tumores (ver **Figura R38b**), decidimos iniciar una nueva línea de investigación sobre el papel de DNMT2 en el contexto del cáncer.

En primer lugar, analizamos las variaciones en el número de copias (CNVs, por sus siglas en inglés), así como el porcentaje de mutaciones puntuales del gen *DNMT2* en pacientes que padecen diversos tipos de cáncer (**Figura R39**). Para llevar a cabo este análisis, empleamos el portal GDC (*National Cancer Institute Genomic Data Commons*) con acceso a datos del Atlas del Genoma del Cáncer (TCGA), que incluye la caracterización molecular de más de 20.000 muestras de tumores y muestras normales asociadas, abarcando 33 tipos de cáncer.

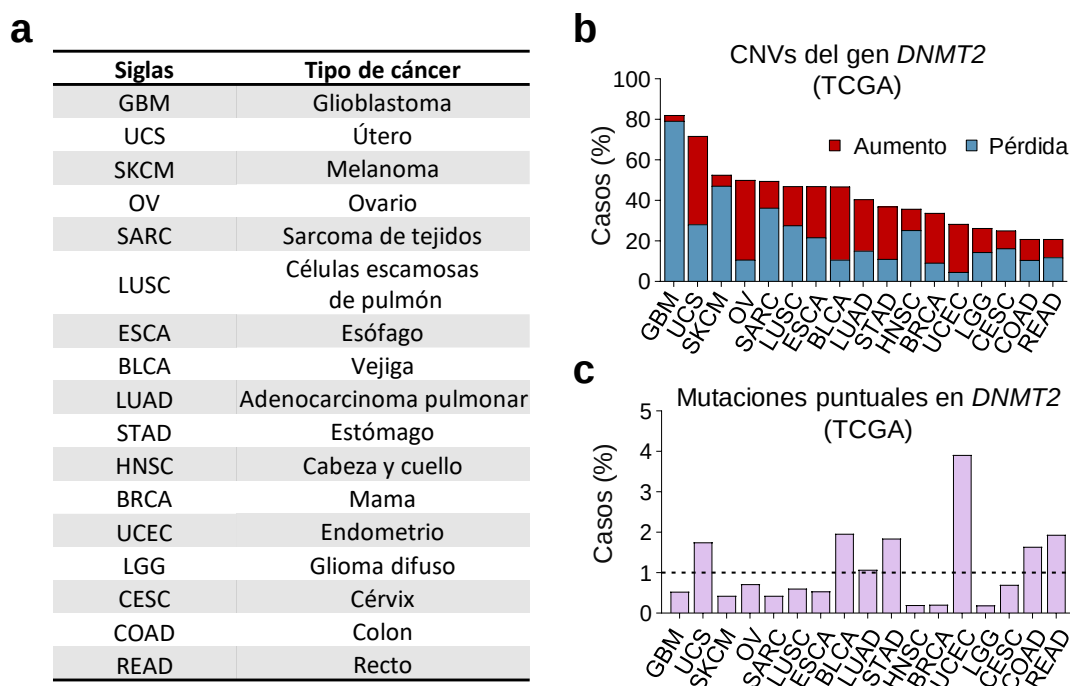


Figura R39. El gen *DNMT2* se encuentra mutado en una amplia variedad de tipos de cáncer. **a)** Tabla en la que se indican los tipos de cáncer en los que al menos el 20% de los pacientes presentan mutaciones en el gen *DNMT2*. En ellos se analizó el porcentaje de CNVs (**b**) y mutaciones puntuales (**c**) del gen *DNMT2*. Se emplearon datos del programa TCGA obtenidos a través del portal GDC. Representación mediante gráfico de barras del porcentaje y tipo de CNVs (**b**) y porcentaje de casos en los que se han detectado mutaciones puntuales (**c**) en los tipos de cáncer indicados.

Nuestro análisis reveló la presencia de mutaciones de tipo CNV en la mayor parte de los tipos de cáncer analizados, siendo en 17 de estos el porcentaje de pacientes afectados superior al 20% (**Figuras R39a y R39b**). Más concretamente, observamos una tendencia marcada hacia la pérdida de copias del gen *DNMT2* en pacientes con glioblastoma (GLB) y melanoma (SKCM), así como una ganancia de copias en aquellos afectados por cáncer de ovario (OV) o de vejiga (BLCA), entre otros (**Figura R39b**). En los otros tipos de cáncer examinados, se observó una proporción similar de casos con pérdida o ganancia de copias en el gen *DNMT2* (**Figura R39b**). No obstante, es importante destacar que el número de mutaciones puntuales detectadas en el gen *DNMT2* es relativamente bajo en la mayoría de los tipos de cáncer analizados (<1%) (**Figura R39c**), encontrándose el mayor número de mutaciones (~4%) en pacientes con cáncer de endometrio (UCEC) (**Figura R39c**).

Posteriormente, llevamos a cabo un análisis más exhaustivo para investigar qué otros genes presentaban mutaciones concurrentes con *DNMT2*, con el objetivo de obtener pruebas sobre posibles mecanismos moleculares que podrían explicar cómo mutaciones en *DNMT2* participan en eventos de transformación oncogénica en pacientes con cáncer. Los resultados revelaron que *TP53* es el gen más frecuentemente mutado en conjunción con *DNMT2* (**Figuras R40a y R40b**) en pacientes con cáncer, seguido de otros genes ampliamente relacionados con esta enfermedad como *TTN* (Wang et al., 2020b) o *MUC16* (Felder et al., 2014) (**Figura R40b**). De hecho, la frecuencia de mutación de estos se ve incrementada cuando *DNMT2* también lo está (**Figura R40b**).

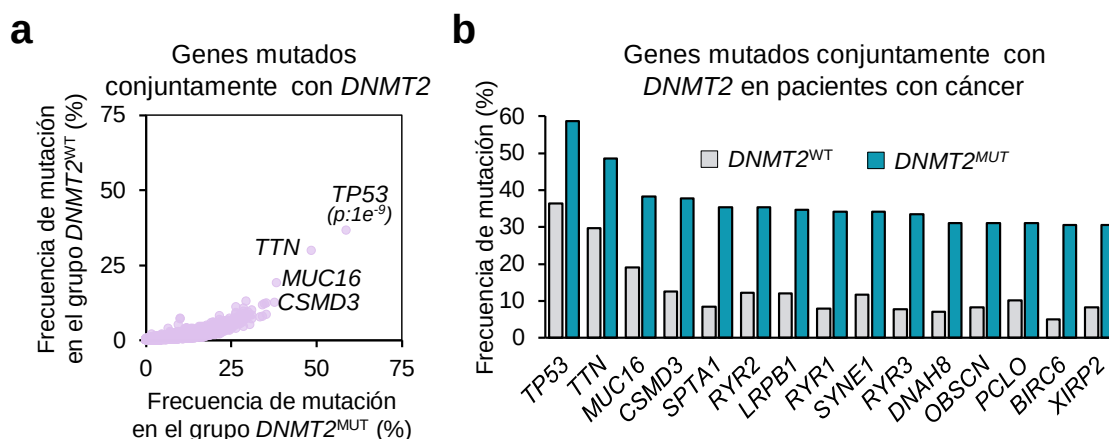


Figura R40. Análisis de los genes con mutaciones conjuntas a *DNMT2* en pacientes con cáncer. a) Representación de los genes más frecuentemente mutados junto con *DNMT2* en pacientes con cáncer. b) Frecuencia de mutación de los genes comúnmente mutados junto con *DNMT2* en pacientes de cáncer, diferenciados por la presencia (azul) o ausencia (gris) de mutaciones en *DNMT2*. Se utilizaron datos del programa TCGA, que incluyen información de 172 pacientes con cáncer con mutaciones en *DNMT2* (*DNMT2*^{MUT}) y 10.788 pacientes con cáncer sin mutaciones en este gen (*DNMT2*^{WT}).

Dado que se ha establecido previamente que las mutaciones en el gen *TP53* se asocian con pronósticos desfavorables en varios tipos de cáncer (Lyu et al., 2019; Robles y Harris, 2010; Ungerleider et al., 2018; Zhao et al., 2020), nos planteamos si las mutaciones en *DNMT2*, así como las alteraciones en sus niveles de expresión, también influyen en la supervivencia de los pacientes. Para abordar esta pregunta, realizamos dos análisis de supervivencia. Por un lado, comparamos la supervivencia de pacientes con mutaciones en *DNMT2* (*DNMT2*^{MUT}) con la de aquellos sin mutaciones en este gen (*DNMT2*^{WT}) (**Figura R41a**), para lo que se emplearon datos del programa TCGA. Por otro lado, se comparó la supervivencia de pacientes con niveles elevados de expresión de *DNMT2* con la de pacientes con niveles reducidos (**Figura R41b**). En

este caso, se emplearon datos de RNA-seq de distintos tipos de cáncer pertenecientes al consorcio Pan Cancer, obtenidos de la base de datos del programa KMPlotter (Nagy et al., 2021). Los resultados obtenidos indicaron que tanto la presencia de mutaciones en *DNMT2* como la disminución de sus niveles de expresión impactan negativamente en la supervivencia de pacientes con cáncer, con cocientes de riesgo calculados de 1,10 y 1,13, respectivamente (Figuras R41a y R41b).

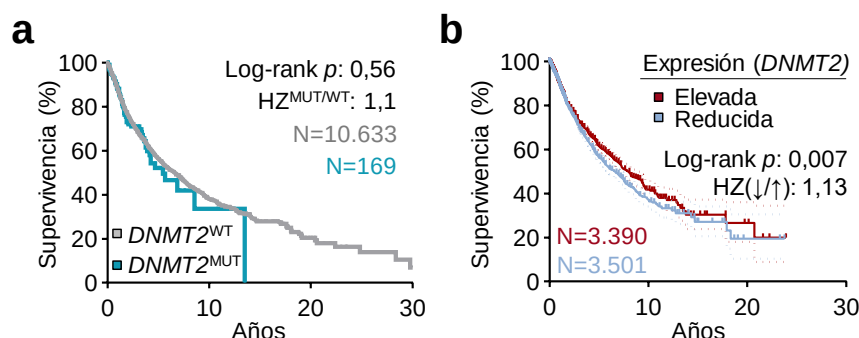


Figura R41. Asociación entre las mutaciones o la reducción de los niveles de *DNMT2* y el pronóstico desfavorable en pacientes con cáncer. a) Análisis de supervivencia Kaplan-Meier que revela la correlación entre la presencia (azul) o ausencia (gris) de mutaciones en *DNMT2* y la supervivencia de los pacientes con cáncer. Se analizaron datos de 10.802 pacientes con información clínica disponible en el repositorio TCGA. b) Análisis de supervivencia Kaplan-Meier que muestra la relación entre la expresión de *DNMT2* y la supervivencia de los pacientes con cáncer. Se emplearon datos de RNA-seq de tumores de 6.891 pacientes pertenecientes al consorcio Pan Cancer. En a y b, los valores *p* y el cociente de riesgo (Hazard Ratio o HR) se calcularon mediante una prueba Mantel-Cox (Log-rank test).

6.3.2 El silenciamiento de *DNMT2* aumenta la agresividad de la línea de cáncer de mama MCF-7

Considerando el impacto de las mutaciones y las alteraciones en la expresión de *DNMT2* en la supervivencia de pacientes con cáncer, nuestro objetivo fue determinar en qué tipos de cáncer la disfunción de *DNMT2* podría tener una mayor relevancia patológica. Para abordar esta cuestión, realizamos un análisis de la frecuencia de mutaciones en *DNMT2* en diferentes tejidos tumorales utilizando datos disponibles en el portal GDC (Figura R42a). El resultado reveló que más del 50% de las mutaciones en *DNMT2* en pacientes con cáncer se encuentran en los tejidos de ovario (16%), mama (14%), pulmón (12%) y útero (11%) (Figura R42a).

Además, tras consultar las bases de datos GTEx y OncoDB (Tang et al., 2022) observamos que en condiciones normales estos tejidos presentan elevados niveles de expresión de *DNMT2* respecto a otros (Figura R42b), regulándose significativamente a la baja en los correspondientes tejidos tumorales (Figura R42c). Coincidiendo con estas observaciones, encontramos que las líneas celulares de cáncer derivadas de estos tumores tienden a mostrar una regulación a la baja en la expresión de *DNMT2* en comparación con las líneas derivadas de otros tipos de cáncer, destacando especialmente las derivadas de tumores mamarios (Figura R42d).

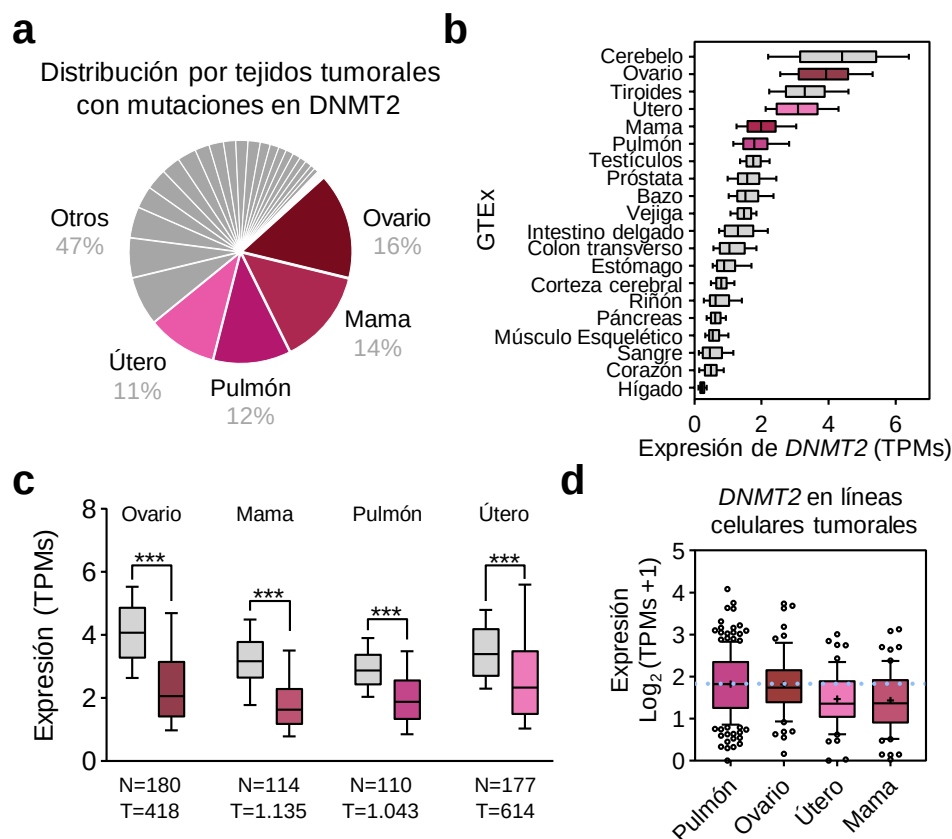


Figura R42. Evaluación de la expresión de DNMT2 en diversos tipos de cáncer. a) Distribución de las mutaciones en el gen *DNMT2* en distintos tipos de tumores. Se analizaron datos de 595 pacientes con información clínica disponible en el repositorio del programa TCGA. b) Niveles de expresión de *DNMT2* en tejidos humanos adultos. Se destacan los tejidos cuyos tumores derivados presentan mutaciones en *DNMT2* con mayor frecuencia (ovario, mama, pulmón y útero). Los datos fueron obtenidos del portal GTEx. c) Distribución de los niveles de expresión obtenidos por RNA-seq de *DNMT2* en tejidos normales (gris) y en los tumores derivados (color). Se muestran los valores de expresión como TPMs o transcritos por millón. Los datos empleados fueron obtenidos de la base de datos OncoDB. Se indica el número de muestras de tejidos normales (N) y tumorales (T) empleadas en el análisis. El cálculo de las diferencias estadísticas se llevó a cabo mediante una prueba Mann Whitney de dos colas no pareada, indicadas como ***: $p < 0,001$. d) Niveles de expresión de *DNMT2* (RNA-seq) en distintas líneas celulares derivadas de los tumores de pulmón (185), de ovario (65), de útero (58) y de mama (62). Los datos, mostrados como transcritos por millón (TPMs), se obtuvieron en el portal DepMap. La línea de puntos azul indica la expresión media de *DNMT2* en 1.450 líneas celulares tumorales.

Dado que los niveles de expresión de *DNMT2* se encuentran significativamente reducidos en tumores mamarios, procedimos a analizar la repercusión de los cambios en su expresión en la supervivencia de los pacientes con cáncer de mama. Para llevar a cabo este análisis, utilizamos datos de RNA-seq de pacientes con cáncer de mama, los cuales contenían información clínica relevante, disponibles en la base de datos del programa KMplotter (Gyorffy, 2021), que reúne datos de 7.830 pacientes pertenecientes a 55 conjuntos de datos diferentes. Dividimos a los pacientes en grupos con elevados o bajos niveles de expresión de *DNMT2* y realizamos un análisis de supervivencia utilizando el programa KMplotter (Gyorffy, 2021). El análisis reveló una reducción significativa de la supervivencia en pacientes con bajos niveles de expresión de *DNMT2* respecto a pacientes con mayores niveles, con un cociente de riesgo de 1,73 (**Figura R43a**). Luego, decidimos identificar si esta correlación entre los niveles de expresión de *DNMT2* y la supervivencia de los pacientes con cáncer de mama se producía de manera general o únicamente en subtipos específicos del mismo.

En el cáncer de mama se conocen clásicamente cuatro subtipos principales en base a sus características moleculares: luminal A, luminal B, HER2+ y triple negativo (Harbeck et al., 2019; Perou et al., 2000) (**Figura R43b**). El subtipo luminal A se caracteriza por altos niveles de los receptores de estrógenos (ER) y progesterona (PR), y bajos niveles del receptor 2 del factor de crecimiento epidérmico humano (HER2), siendo su perfil inmunológico ER+ PR+ HER2-. De manera similar, el subtipo luminal B presenta un perfil inmunológico ER+ PR- HER2+, aunque también se integran en este grupo aquellos tumores con perfil ER+ PR- HER2. Por otro lado, el subtipo HER2+ se caracteriza por elevados niveles del receptor HER2 y bajos o nulos niveles de los receptores ER y PR, mientras que el triple negativo carece de los receptores ER y PR y presenta niveles muy bajos o nulos de HER2 (Harbeck et al., 2019).

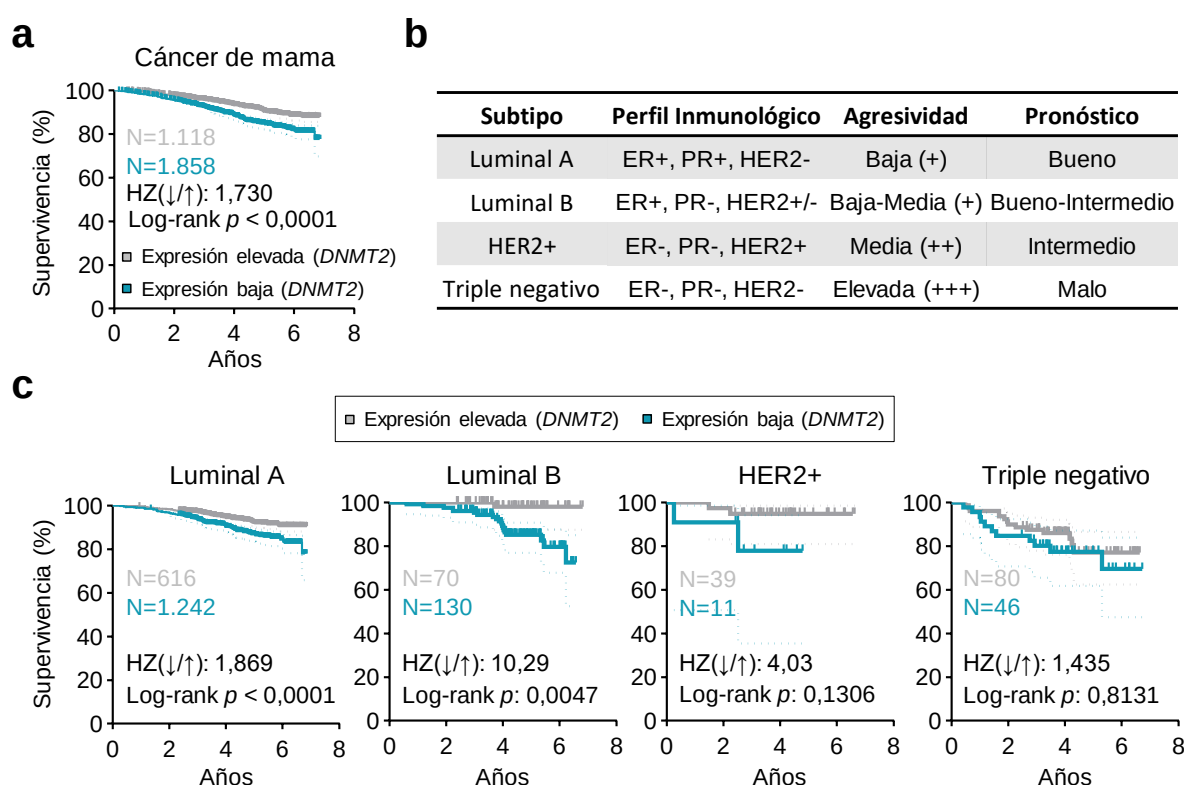


Figura R43. Asociación entre bajos niveles de *DNMT2* y una menor supervivencia en pacientes con cáncer de mama de los subtipos luminal A y luminal B. a) Análisis Kaplan-Meier que muestra la correlación entre bajos niveles de expresión del gen *DNMT2* y una menor supervivencia de pacientes con cáncer de mama. Se emplearon datos de 2.976 pacientes que se incluyen dentro de los 4 subtipos principales de cáncer de mama. b) Tabla resumen que describe los subtipos de cáncer de mama analizados mediante las variables: perfil inmunológico, agresividad y pronóstico. c) Análisis Kaplan-Meier de supervivencia que muestra como los niveles de expresión de *DNMT2* correlacionan con la supervivencia de pacientes con cada uno de los distintos subtipos de cáncer de mama. En a y c el *p* valor y el cociente HZ fueron calculados mediante una prueba Mantel-Cox (*Log-rank*). Se emplearon datos de RNA-seq de muestras de pacientes de cáncer de mama disponibles en la base de datos del programa KMPlotter (Gyorffy, 2021).

Los subtipos luminales A y B se consideran poco agresivos y presentan un pronóstico relativamente bueno, ya que responden con eficacia a tratamientos antitumorales (Harbeck et al., 2019). Estas características se ven agravadas en el caso del subtipo HER2+, caracterizado por una mayor agresividad y un peor pronóstico (Harbeck et al., 2019), mientras que el triple negativo exhibe una agresividad elevada y un mal pronóstico debido a la ausencia de proteínas empleadas como diana en la mayoría de las terapias antitumorales, lo que dificulta su tratamiento (Harbeck et al., 2019) (**Figura R43b**). En este análisis, se detectó una reducción

significativa de la supervivencia de los pacientes con subtipos luminal A y luminal B ante bajos niveles de expresión de *DNMT2*, no observándose cambios significativos en la supervivencia de los pacientes con los subtipos HER2⁺ o triple negativo (**Figura R43c**). En base a esto, la reducción de los niveles de expresión de *DNMT2* correlaciona con un empeoramiento del pronóstico en pacientes con tumores poco agresivos, sin afectar a subtipos tumorales de mayor agresividad.

Hasta el momento, nuestros resultados sugieren una posible correlación entre las mutaciones en los genes *DNMT2* y *TP53* (**Figura R40**), así como la disminución de los niveles de expresión de *DNMT2* y un pronóstico desfavorable en pacientes con cáncer de mama de los subtipos menos agresivos (luminales A y B) (**Figura R43c**). Considerando esta información, decidimos investigar en un entorno de cultivo celular si la pérdida de función de *DNMT2* puede potenciar el fenotipo agresivo de células tumorales con baja capacidad invasiva. Para llevar a cabo este análisis, optamos por utilizar las líneas tumorales humanas MCF-7 y MDA-MB-231, proporcionadas por nuestro colaborador, el Dr. Román Pérez en el CIMUS. Estas líneas representan subtipos de tumores poco y muy invasivos, respectivamente, y difieren en el estado mutacional del gen *TP53* (**Figura R44a**) (Harbeck et al., 2019).

La línea MCF-7 está asociada con el subtipo luminal A, caracterizado por niveles elevados de ER y PR, y bajos niveles de HER2 (Dai et al., 2017; Harbeck et al., 2019). En cambio, la línea celular MDA-MB-231 pertenece al subtipo triple negativo con un perfil inmunológico ER- PR- HER2- y presenta una alta capacidad invasiva (Dai et al., 2017; Harbeck et al., 2019). Además de las diferencias en el perfil inmunológico, la capacidad invasiva y otras características, MCF-7 presenta una versión WT de la proteína P53 (Berglind et al., 2008), mientras que se han identificado distintas mutaciones en el gen que codifica P53 en las células MDA-MB-231 (Adorno et al., 2009; Girardini et al., 2011). Estas mutaciones, en algunos casos, contribuyen a la supervivencia de las células tumorales (Adorno et al., 2009; Girardini et al., 2011; Muller y Vousden, 2014).

Para poner a prueba nuestra hipótesis de que la reducción de los niveles de *DNMT2* podría agravar el fenotipo del subtipo tumoral menos agresivo, realizamos ensayos de pérdida de función utilizando shRNAs dirigidos contra el gen humano *DNMT2*, con sh*Luci* como control, en las líneas MCF-7 y MDA-MB-231. Luego, evaluamos las capacidades de proliferación, formación de colonias y migración en cada condición (**Figura R44b**).

En primer lugar, confirmamos la reducción significativa de los niveles de expresión de *DNMT2* en ambas líneas tumorales mediante RT-qPCR después de la transducción con dos shRNAs independientes (**Figura R44c**). Posteriormente, realizamos un ensayo de formación de colonias (CFA) en agar blando con el fin de evaluar la capacidad de crecimiento independiente de anclaje de las células tumorales (Mori et al., 2009). Los resultados mostraron un aumento significativo en la capacidad de formación de colonias independientemente de anclaje, solo en la línea MCF-7 y no en MDA-MB-231, después de la inhibición de *DNMT2* en comparación con el control sh*Luci* (**Figura R44d**). Por otro lado, no se observaron diferencias significativas en la capacidad de proliferación de las células MCF-7 o MDA-MB-231 después de la inhibición de *DNMT2* (**Figura R44e**).

Estos resultados muestran un incremento en la capacidad de formación de colonias de células tumorales poco agresivas, independientemente de su capacidad de proliferación. Sin embargo, no se observó este efecto en células con un fenotipo más agresivo.

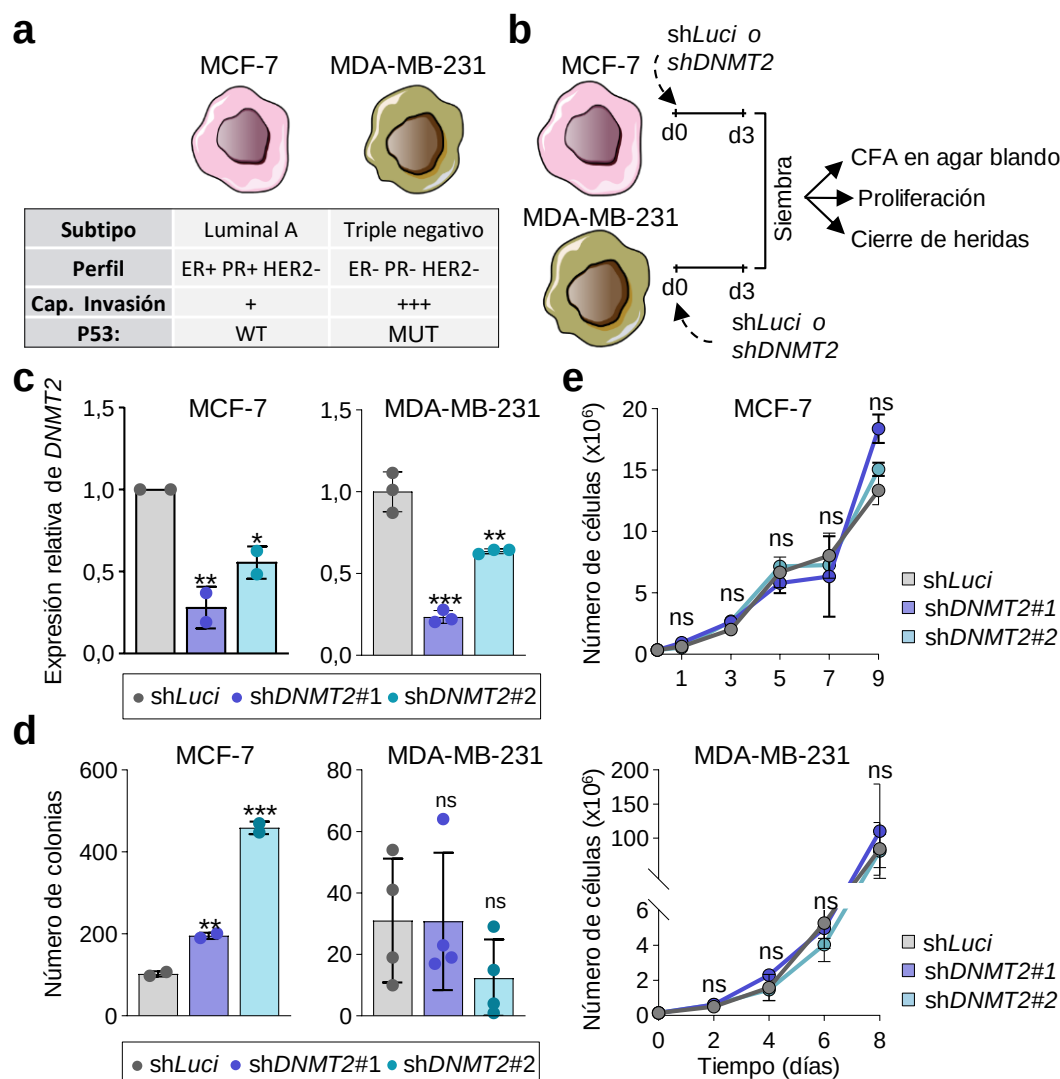


Figura R44. La inhibición de DNMT2 incrementa la capacidad de formación de colonias en la línea tumoral MCF-7. a) Descripción de las líneas de cáncer de mama MCF-7 y MDA-MB-231 en base al subtipo tumoral al que pertenecen, su perfil inmunológico, su capacidad de invasión y el estado de la proteína P53. b) Estrategia seguida para evaluar el efecto de la pérdida de función de DNMT2 sobre la capacidad de formación de colonias, proliferación y migración, de células MCF-7 o MDA-MB-231, mediante ensayos de formación de colonias (CFA) en agar blando, proliferación y cierre de heridas, respectivamente. c) Análisis de los niveles de expresión de hDNMT2, en células MCF-7 (izquierda) o MDA-MB-231 (derecha), transducidas con shRNAs contra DNMT2 o shLuci, por RT-qPCR. Se muestra la expresión relativa a U6 y a shLuci. d) Número de colonias en agar blando obtenidas a partir de células MCF-7 (izquierda) o MDA-MB-231 (derecha), transducidas con los shRNAs indicados. e) Representación del recuento del número de células MCF-7 (panel superior) o MDA-MB-231 (panel inferior) transducidas con los shRNAs indicados, en los días especificados. En c, d y e, los datos se presentan como media $\pm$ desviación estándar, indicándose el número de replicados biológicos mediante puntos. Para el cálculo de diferencias estadísticamente significativas se aplicó una prueba ANOVA de un factor no pareada (c y d), o una prueba ANOVA de dos factores pareada (e), indicadas como ns: no significativo, *: $p < 0,05$, **: $p < 0,01$ y ***: $p < 0,001$.

Finalmente, evaluamos la capacidad de migración de las células MCF-7 y MDA-MB-231 después de la pérdida de función de DNMT2. Para esto, llevamos a cabo un ensayo de cierre de heridas durante 36 horas para las células MCF-7 y 15 horas para las células MDA-MB-231. De manera significativa, observamos que las células MCF-7 transducidas con shRNAs contra DNMT2 presentaban una mayor capacidad de migración, lo que resultó en un cierre más rápido de las heridas formadas (**Figuras R45a y R45b**). Por lo contrario, no se

detectaron diferencias relevantes en la capacidad de migración de células MDA-MB-231 tras la reducción de los niveles de expresión de *DNMT2* (Figuras R45c y R45d).

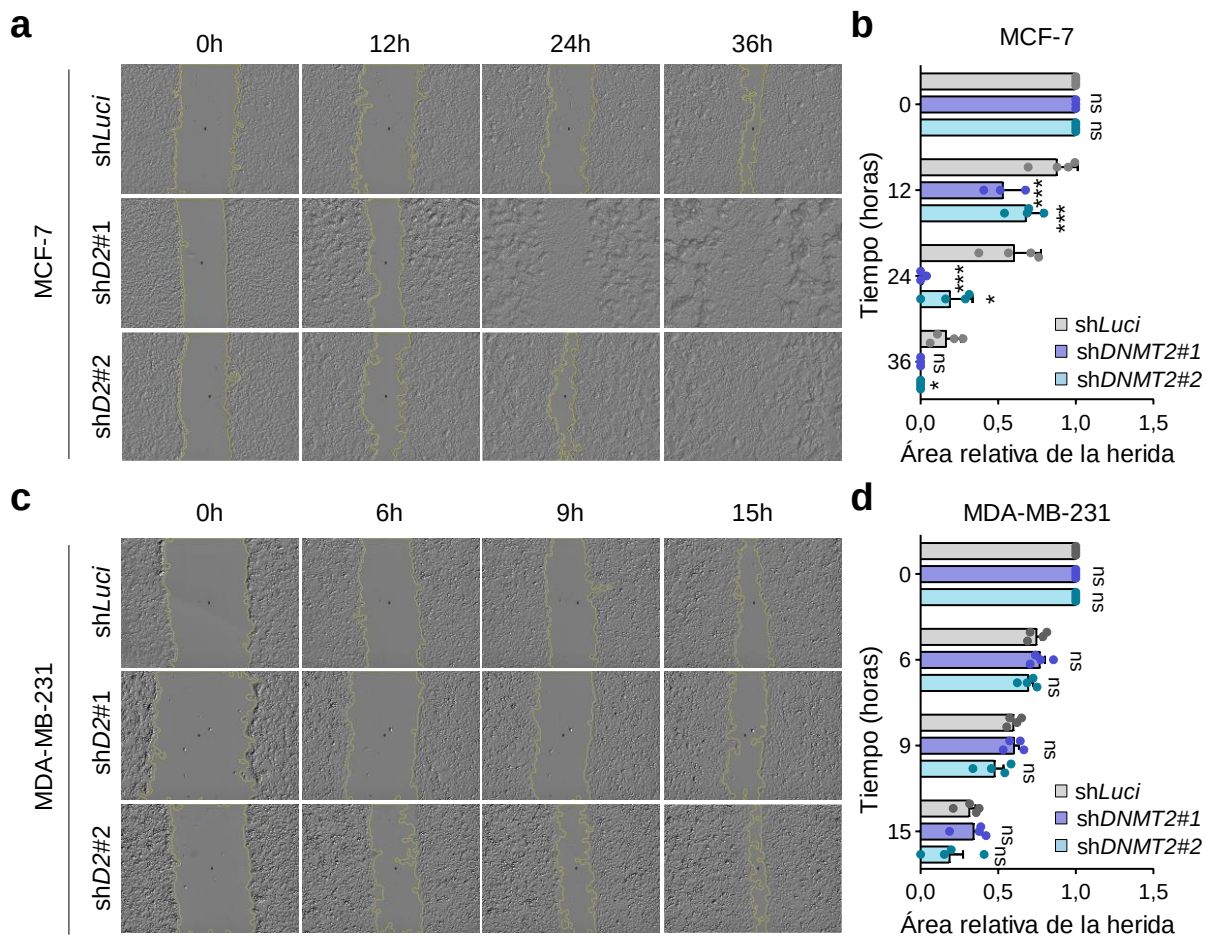


Figura R45. El silenciamiento de *DNMT2* incrementa la capacidad de migración en la línea tumoral MCF-7. a) Imágenes representativas del ensayo de cierre de heridas en células MCF-7 transducidas con shRNAs contra *DNMT2* o *shLuci*, en intervalos de 0, 12, 24 y 36 horas desde el trazado de la herida. Se señala en amarillo el área de la herida analizada. b) Cuantificación del área de la herida en células MCF-7 transducidas con los shRNAs indicados, relativa al día 0, en los tiempos mencionados. c) Imágenes representativas del ensayo de cierre de heridas en células MDA-MB-231 transducidas con los shRNAs indicados, en intervalos de 0, 6, 9 y 15 horas desde el trazado de la herida. Se señala en amarillo el área de la herida analizada. d) Cuantificación del área de la herida en células MDA-MB-231 transducidas con los shRNAs indicados, relativa al día 0, en los tiempos mencionados. shD2: *shDNMT2*. En b y d, los datos se representan como media $\pm$ desviación estándar y los replicados biológicos se indican mediante puntos. Se utilizó una prueba ANOVA de dos factores pareada para calcular las diferencias estadísticamente significativas, indicadas como ns: no significativo, *: $p < 0,05$, **: $p < 0,01$ y ***: $p < 0,001$.

Por lo tanto, los datos obtenidos apoyan la hipótesis planteada previamente, sugiriendo que la pérdida de función de *DNMT2* provoca un agravamiento del fenotipo de células tumorales poco agresivas, sin que esto ocurra en células con un fenotipo más agresivo.

6.3.3 *DNMT2* participa en la regulación de P53 en la línea tumoral MCF-7

Nuestros hallazgos en ESCs de origen murino indican que la eliminación de *DNMT2* resulta en un aumento de los niveles de expresión de *Mdm2* y una disminución de los niveles de la proteína P53 (Figura R23). Además, posteriores análisis *in silico* revelaron que *TP53* es el gen más mutado en conjunto con *DNMT2* en pacientes con cáncer (Figura R40).

Teniendo esto en consideración y el hecho de que el silenciamiento de *DNMT2* conduce a un fenotipo más agresivo en las células MCF-7, que poseen una forma funcional de P53 (P53^{WT}), pero no en las células MDA-MB-231 con una variante mutada de P53 (P53^{MUT}) (**Figuras R44 y R45**), nos planteamos la posibilidad de que *DNMT2* podría intervenir en la regulación de la abundancia de P53 en la línea tumoral MCF-7, la cual es representativa del subtipo menos agresivo de cáncer de mama (luminal A).

Con el propósito de abordar esta cuestión, se transdujeron células MCF-7 con shRNAs contra *DNMT2* o *Luciferasa* como control. Luego, analizamos los niveles de expresión de *MDM2* mediante RT-qPCR y la abundancia de P53 mediante WB (**Figura R46a**). En consonancia con los resultados previos obtenidos en ESCs, observamos un aumento significativo en la expresión de *MDM2* después del silenciamiento de *DNMT2* en células MCF-7 (**Figura R46b**). Además, se apreció una correlación negativa entre los niveles de expresión de *DNMT2* y los de *MDM2*, con un coeficiente de correlación del 53,94% en esta línea tumoral (**Figura R46c**).

Dado que *MDM2* es el principal regulador negativo de P53 (Biegging et al., 2014; Marine y Lozano, 2010), sería de esperar una reducción en la abundancia de los niveles proteicos de P53 en estas mismas condiciones. De acuerdo con esta posibilidad, el análisis por WB reveló una disminución en los niveles de la proteína P53 después del silenciamiento de *DNMT2*, en comparación con las células transducidas con sh*Luci* (**Figuras R46d y R46e**).

Lo siguiente que nos propusimos, fue determinar si la actividad m⁵C-metiltransferasa de *DNMT2* estaba involucrada en el control de la expresión del gen *MDM2* en estas células. Para ello, procedimos a transducir las células MCF-7 con el sistema de sobreexpresión lentiviral FUW-Tet-On que contenía la versión humana WT de *DNMT2* (h*DNMT2*^{WT}) o su versión mutante inactiva (h*DNMT2*^{C79A}) (**Figura R46f**). Tras tres días de exposición a doxiciplina (DOX) en presencia del transactivador M2rtTA, confirmamos la sobreexpresión de la forma WT y de la mutante (C79A) mediante RT-qPCR (**Figura R46g**). Bajo estas mismas condiciones, se observó una reducción significativa de los niveles de expresión de *MDM2* tras la sobreexpresión de la proteína catalíticamente activa (WT), pero no de la mutante catalítica (C79A) (**Figura R46h**).

Los datos obtenidos sugieren que *DNMT2* participa en la regulación transcripcional de *MDM2* a través de su actividad catalítica en células de cáncer de mama del subtipo luminal A. En consecuencia, cabría esperar que los niveles de la proteína P53 en este tipo de células también dependiesen de esta actividad. No obstante, es importante destacar que los resultados obtenidos no excluyen la posibilidad de que *DNMT2* regule los niveles de P53 a través de su función no catalítica. Vale la pena señalar que uno de los principales interactores de *DNMT2* identificados en ESCs, la proteína 14-3-3E (**Figura R36**), pertenece a una familia de proteínas implicadas en la regulación transcripcional de P53, así como en su estabilidad (Falcicchio et al., 2020).

En resumen, nuestros resultados sugieren que el aumento en el fenotipo más agresivo observado en la línea celular MCF-7 después de la pérdida de función de *DNMT2* podría estar relacionado con una regulación al alza del gen *MDM2* y la consecuente reducción de los niveles de la proteína P53. No obstante, se requerirán futuros estudios para validar esta posibilidad o explorar otros posibles mecanismos moleculares subyacentes.

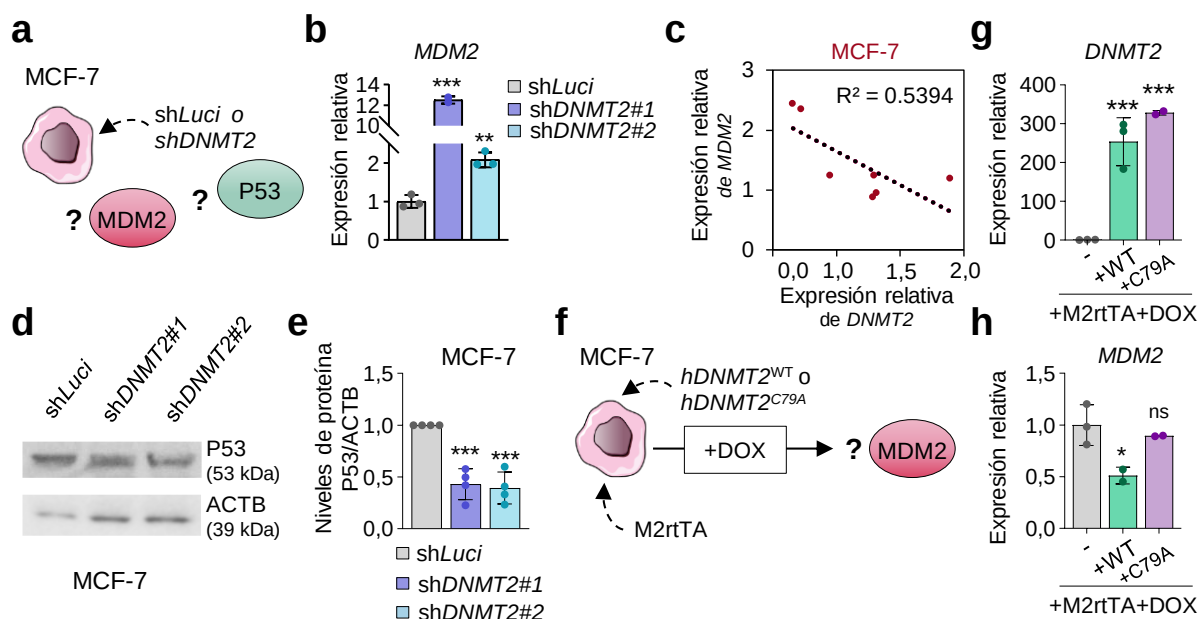


Figura R46. La inhibición de *DNMT2* resulta en una disminución de los niveles de la proteína P53 en la línea tumoral MCF-7. **a)** Esquema de la estrategia seguida. las células MCF-7 fueron transducidas con shRNAs contra *DNMT2* o *Luciferasa* como control, seguido del análisis de los niveles de *MDM2* y P53 mediante RT-qPCR y WB, respectivamente. **b)** Análisis por RT-qPCR de los niveles de *MDM2* en las condiciones indicadas. Se muestra la expresión relativa a los niveles de *U6* y a *shLuci*. **c)** Correlación entre los niveles relativos de expresión de *DNMT2* y *MDM2* en células MCF-7, indicando el coeficiente de correlación (R^2). **d)** Niveles proteicos de P53 en células MCF-7 después de la inhibición de *DNMT2*, evaluados por WB. Se presentan los niveles de ACTB como control de carga. **e)** Cuantificación de los niveles de la proteína P53 normalizados al control de carga ACTB en células MCF-7 tras el silenciamiento de *DNMT2*. **f)** Esquema de la estrategia seguida para determinar el efecto de la sobreexpresión *DNMT2* WT o mutante catalítico (C79A) en la regulación transcripcional de *MDM2* en células MCF-7. **g)** Análisis por RT-qPCR de los niveles de *DNMT2* en las condiciones indicadas. **h)** Análisis por RT-qPCR de los niveles de *MDM2* en las condiciones indicadas. En g y h, se presenta la expresión relativa respecto a los niveles de *U6* y células tratadas con DOX en ausencia de las formas sobreexpresadas de *DNMT2*. En b, e, g y h, los datos se muestran como media $\pm$ desviación estándar, indicándose por puntos el número de réplicas biológicas. Para la evaluación de diferencias estadísticamente significativas se llevó a cabo una prueba ANOVA de un factor no pareada, con corrección de Dunnet para comparaciones múltiples, indicadas como ns: no significativo, *: $p < 0,05$, **: $p < 0,01$ y ***: $p < 0,001$.

7. DISCUSIÓN

Aunque las modificaciones químicas del ARN desempeñan un papel crucial en la regulación de la identidad y plasticidad celulares, su desregulación conlleva importantes implicaciones en diversos procesos patológicos. En la actualidad, se está llevando a cabo una intensa investigación sobre las funciones específicas de la marca m^5C y las proteínas que regulan su metabolismo. DNMT2 destaca entre los miembros de la familia de las ADN metiltransferasas (DNMTs), ya que cataliza la deposición de m^5C en moléculas de ARN en mamíferos, en lugar de en ADN. A pesar de su alto grado de conservación evolutiva, la relevancia biológica funcional de DNMT2 sigue sin estar resuelta, ya que los organismos modelo sobre los que se ha estudiado su deficiencia muestran fenotipos divergentes y, en general, bastante sutiles. En el transcurso de esta tesis doctoral se ha llevado a cabo una caracterización detallada del papel de DNMT2 en los contextos de pluripotencia y cáncer.

En el ámbito de pluripotencia, se exploraron las funciones dependientes e independientes de la actividad catalítica de DNMT2 durante la regulación del estado celular pluripotente, abarcando los procesos de adquisición y salida del mismo. En primer lugar, demostramos que la actividad catalítica de DNMT2 es esencial para el establecimiento de programas transcripcionales epiteliales en la transición mesenquimal-epitelial (MET) durante la reprogramación celular somática (SCR) para la obtención de iPSCs, y que su sobreexpresión favorece este proceso. Además, su función también resultó ser clave en eventos de reprogramación celular directa para el establecimiento de otros destinos celulares, como el hematopoyético. Por otro lado, observamos que, aunque la función de DNMT2 es prescindible en el mantenimiento de la pluripotencia en cultivo, su actividad catalítica resulta crítica para la activación transcripcional de genes implicados en la correcta salida del estado celular pluripotente.

Finalmente, en el contexto de cáncer se identificó que las mutaciones en el gen que codifica a DNMT2 o la disminución de su expresión correlacionaban con un pronóstico desfavorable en pacientes afectados por diversos tipos de cáncer, incluyendo el cáncer de mama. Además, se observó que el silenciamiento de DNMT2 en la línea tumoral MCF-7, representativa del subtipo luminal A de cáncer de mama, resultaba en un fenotipo más agresivo de estas células, posiblemente relacionado con la regulación a la baja de los niveles proteicos de P53 debido a la ausencia de DNMT2.

7.1 LA CONTRIBUCIÓN DE DNMT2 A LA REPROGRAMACIÓN CELULAR

En este trabajo revelamos un papel, hasta el momento desconocido, de la proteína DNMT2 en la adquisición de pluripotencia mediante la SCR mediada por OSKM. Aunque diversos estudios han documentado la participación de otras modificaciones del ARN durante este proceso, como la N6-metiladenosina (m^6A) (Aguilo et al., 2015; Chen et al., 2015; Khodeer et al., 2022; Liu et al., 2020b) y la desaminación de adenosina en inosina (A-to-I) (Guallar et al., 2020), la relevancia de nuestros hallazgos radica en que el papel de la modificación m^5C y sus enzimas reguladoras sigue siendo desconocido en este contexto.

La función de DNMT2 ha sido previamente asociada con cambios de identidad celular en eventos de diferenciación en distintos organismos (Ghanbarian et al., 2016; Rai et al., 2007; Tuorto et al., 2015). Por ejemplo, durante el desarrollo embrionario del pez cebra, DNMT2 desempeña un papel crítico en la diferenciación celular que tiene lugar en órganos como el hígado, la retina y el cerebro (Rai et al., 2007). En ratón, su función es necesaria tanto durante la hematopoyesis (Tuorto et al., 2015), como durante la formación de tejido cardíaco (Ghanbarian et al., 2016). No obstante, su contribución en el proceso inverso de reprogramar

células diferenciadas hacia células con un mayor potencial de diferenciación, como es el caso de las células pluripotentes, todavía no ha sido explorado.

Aquí, mostramos que la pérdida de la modificación m^5C catalizada por DNMT2 compromete la obtención de iPSCs a partir de fibroblastos (**Figuras R1-R3**), dificultando la activación y supresión de programas transcripcionales epiteliales y mesenquimales, respectivamente, durante la MET (**Figuras R4-R7**). Además, la eficiencia de este proceso se ve incrementada tras la sobreexpresión de la proteína DNMT2 WT y no de una variante mutante catalítica (C79A) (**Figura R8**). De este modo, demostramos por primera vez la esencialidad de la actividad catalítica de DNMT2 en las etapas iniciales del proceso de SCR, destacando que resulta prescindible una vez que la identidad celular epitelial ha sido establecida (**Figura R9**).

En línea con nuestros resultados, la regulación de otras modificaciones epitranscriptómicas también resulta crítica durante la MET (Guallar et al., 2020; Khodeer et al., 2022; Liu et al., 2020b). Aquí observamos que la deposición de m^5C es fundamental para la activación de genes epiteliales y la represión de genes mesenquimales, mientras que otros estudios han demostrado que la marca m^6A también juega un papel crucial en este proceso (Khodeer et al., 2022; Liu et al., 2020b). Específicamente, las proteínas YTHDF2 y YTHDF3, reconocedoras ("readers") de m^6A , promueven el reclutamiento de complejos deadenilasa implicados en la degradación de transcritos somáticos marcados con esta modificación, como *Tead2* o *Tgfb1*, facilitadores del proceso opuesto, la transición epitelial-mesenquimal (EMT) (Liu et al., 2020b). De este modo, su silenciamiento también compromete el correcto transcurso de la MET (Liu et al., 2020b). En concordancia, la pérdida de función de ALKBH5, "eraser" de m^6A , impide la regulación al alza de transcritos epiteliales durante este proceso (Khodeer et al., 2022). Al igual que en el caso de DNMT2, su eliminación durante etapas tempranas de la SCR reduce significativamente la eficiencia de reprogramación (Khodeer et al., 2022). Además, la pérdida de la función de ADAR1, implicada en la modificación A-to-I, induce respuestas de estrés en el retículo endoplasmático, comprometiendo así el progreso adecuado de la MET (Guallar et al., 2020). En conjunto, estas evidencias sugieren la necesidad de un riguroso equilibrio entre distintas modificaciones del ARN en transcritos específicos, durante las fases tempranas de la adquisición de pluripotencia, a partir de células somáticas de origen mesenquimal.

En relación con DNMT2, es importante destacar que los defectos observados en la SCR tras su pérdida de función no están relacionados con cambios en la tasa de proliferación, ni en las células somáticas a reprogramar, ni tras el inicio de la reprogramación mediante OSKM (**Figura R2**). Contrariamente, algunos estudios documentan defectos en la proliferación de fibroblastos murinos inmortalizados (Lewinska et al., 2017) y humanos, tanto inmortalizados como no inmortalizados (Lewinska et al., 2018) tras el silenciamiento de *DNMT2*. Estas discrepancias con nuestras observaciones pueden atribuirse, en parte, al propio proceso de inmortalización. En ocasiones, las células sometidas a este tipo de procedimientos exhiben atributos ausentes en las mismas células *in vivo*, como la proliferación indefinida y la ausencia de inhibición por contacto, entre otros. Además, en el caso de los fibroblastos humanos, se observa una marcada reducción de su proliferación tras el silenciamiento de *DNMT2* en combinación con condiciones de estrés oxidativo, atenuándose o desapareciendo este efecto en condiciones normales (Lewinska et al., 2018), lo que concuerda con la falta de efecto sobre el crecimiento de los fibroblastos murinos que observamos.

Además, es importante destacar que la actividad catalítica de DNMT2 no solo favorece la generación de iPSCs, sino que también es necesaria para la reprogramación celular directa,

conocida como transdiferenciación, de fibroblastos a células precursoras hematopoyéticas inducidas (iHSCs) (**Figuras R10-R13**). Este papel de DNMT2 no se limita a células murinas, puesto que también se observa en células humanas (**Figuras R10-R13**), indicando que se trata de una función conservada entre organismos. Nuestros resultados, junto con evidencias previas que demuestran su papel catalítico-dependiente en la diferenciación durante la hematopoyesis en ratones (Tuorto et al., 2015), sugieren que DNMT2 desempeña un papel fundamental en el establecimiento del destino celular hematopoyético mediante su actividad catalítica en moléculas de ARN.

Aunque en este trabajo hemos demostrado la importancia de la función catalítico-dependiente de DNMT2 durante la adquisición de nuevas identidades celulares, incluida la pluripotencia, aún queda por elucidar el mecanismo molecular mediante el cual la metilación de ARN mediada por DNMT2 regula este proceso (**Figura D1**).

Mientras que las moléculas de ARNt que portan Asp, Val y Gly son consideradas los principales sustratos de DNMT2 (Goll et al., 2006; Jurkowski et al., 2008; Legrand et al., 2017; Schaefer et al., 2010; Tuorto et al., 2015; Zhang et al., 2018), varios estudios han mostrado la metilación dependiente de esta proteína en otras especies de ARN, como ARNm en células tumorales (Chen et al., 2020) o ARNnc en ESCs y células cardíacas (Ghanbarian et al., 2016).

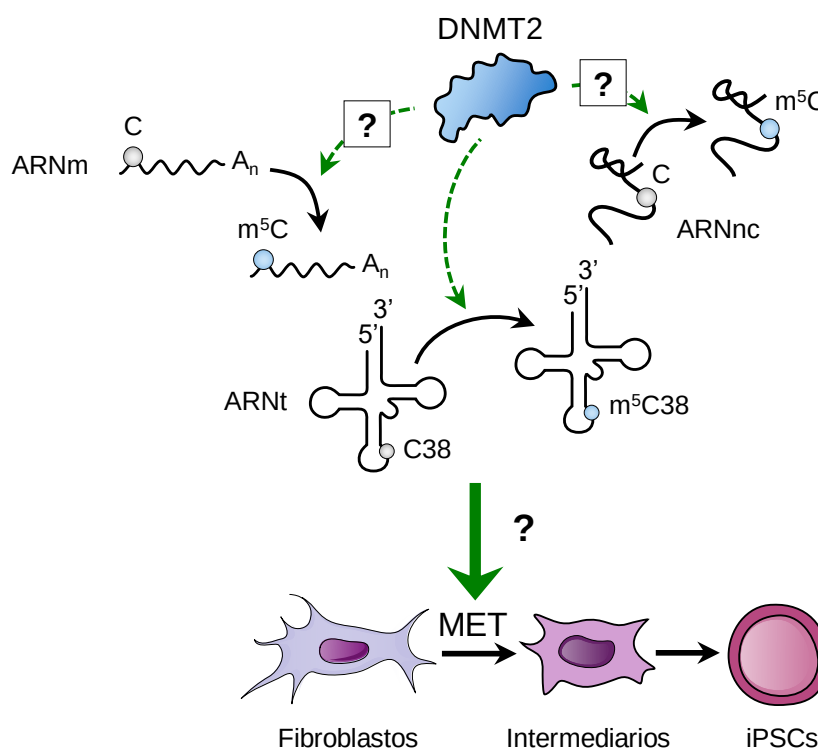


Figura D1. Participación de DNMT2 en la SCR mediante su actividad catalítica. DNMT2 promueve la MET durante la SCR a través de su actividad catalítica sobre el ARN. Sin embargo, desconocemos las moléculas específicas de ARN sobre las cuales DNMT2 cataliza la deposición de m⁵C y que están involucradas directamente en la MET. Las flechas en verde indican el favorecimiento del proceso y los signos de interrogación indican las incógnitas por resolver. Esta imagen fue elaborada mediante el uso del programa Inkscape (v1.3).

Con el fin de discernir si la metilación de ARNt catalizada por DNMT2 es esencial durante la MET, sería interesante analizar cambios en los niveles de estas moléculas, así como de los fragmentos derivados de su escisión (tsRNAs). Estos tsRNAs se encuentran ampliamente

relacionados con la pérdida de m⁵C38 en varios contextos celulares murinos, como células de la médula ósea (Tuorto et al., 2015) o células espermáticas (Yu et al., 2021; Zhang et al., 2018), entre otros (Durdevic et al., 2013b; Schaefer et al., 2010). En este caso, se esperaría un aumento en la generación de tsRNAs como resultado de la pérdida de m⁵C38 después del silenciamiento de *Dnmt2*, atenuándose este efecto tras la expresión forzada de una forma WT, pero no de una variante catalíticamente inactiva de la proteína (C79A). Dada la función crítica de DNMT2 durante la MET, se anticiparían cambios sustanciales en los niveles de tsRNAs en las primeras etapas del proceso de SCR tras su silenciamiento, pero no una vez establecido el destino celular epitelial.

Aunque no se ha explorado en profundidad durante la reprogramación hacia células pluripotentes, los tsRNAs pueden modular la expresión génica en otros contextos. Esto puede ocurrir a través de su interacción con proteínas de unión al ARN que estén implicadas en la estabilización de transcritos, como la proteína YBX1 en células de cáncer de mama (Falconi et al., 2019; Goodarzi et al., 2015), o en su silenciamiento, como las proteínas ARGONAUTA, cuya asociación con este tipo de moléculas se ha descrito en células humanas (Kumar et al., 2014; Kuscu et al., 2018). Por lo tanto, a pesar de que los mecanismos descritos ocurren en contextos celulares diferentes, teniendo en cuenta el elevado grado de conservación de DNMT2 entre organismos muy diversos (Jeltsch et al., 2017; Jurkowski y Jeltsch, 2011), no sería sorprendente que esta proteína también esté implicada en la biogénesis de tsRNAs durante la reprogramación celular.

Por otro lado, en base a nuestros resultados, no podemos descartar la posibilidad de que DNMT2 actúe directamente sobre otros tipos de ARN, como ARNnc o retrotransposones. Esta posibilidad ha sido documentada en células pluripotentes, donde se observó una reducción significativa en la metilación del ARNnc *Rn7SK* en ESCs KO para *Dnmt2* (Ghanbarian et al., 2016). Estos potenciales sustratos podrían encontrarse directamente implicados en la MET, o actuar regulando a otras moléculas que si lo están. De hecho, se ha observado una regulación al alza de ciertos retrotransposones durante la obtención de iPSCs a partir de células somáticas murinas (Friedli et al., 2014) o humanas (Klawitter et al., 2016; Wissing et al., 2012). Por lo tanto, se requiere una investigación más profunda para delinear el mecanismo molecular por el cual DNMT2 promueve la MET en las etapas tempranas de la SCR, así como otros cambios en la identidad celular.

Es importante destacar que, aunque en este trabajo demostramos que la actividad catalítica de DNMT2 es fundamental durante eventos de reprogramación celular, no podemos descartar su participación simultánea en estos procesos a través de una función catalítico-independiente. Un ejemplo de esto ocurre en el caso de la marca m⁶A, cuya regulación es crítica durante la SCR, ya que algunas enzimas que colaboran en su deposición, como METTL14, parte del complejo catalizador de m⁶A junto con METTL3, entre otras proteínas, pueden promover la SCR de manera independiente a esta modificación epitranscriptómica (Xi et al., 2022). Además, según revelan nuestros resultados obtenidos en ESCs, DNMT2 interacciona con proteínas involucradas en el procesamiento del ARN (**Figura R38a**). Teniendo esto en cuenta, y considerando que los defectos observados en la MET ocurren a nivel transcripcional, el silenciamiento de DNMT2 podría alterar la regulación de transcritos implicados en este proceso. De hecho, la inhibición del gen que codifica una de las proteínas identificadas como interactora de DNMT2, BASP1, promueve la generación de iPSCs (Blanchard et al., 2017). Sin embargo, la interacción de DNMT2 con otras proteínas puede variar entre ESCs y las células que experimentan el proceso de SCR, por lo que, con el fin de dilucidar una potencial función

no catalítica, sería interesante analizar con qué proteínas interacciona en distintos puntos de este proceso, mediante espectrometría de masas, por ejemplo. De este modo, las posibles funciones catalítico-independientes de DNMT2 constituyen un área de investigación activa que requiere el diseño de estudios dirigidos a aclarar su existencia.

7.2 LA RELEVANCIA DE DNMT2 EN PLURIPOTENCIA

Las características inherentes de autorrenovación y diferenciación exhibidas por las células pluripotentes en cultivo, ya sean iPSCs o ESCs, las erigen como un modelo celular excepcional para la investigación de los fundamentos moleculares que regulan la plasticidad celular. Esta importancia se enfatiza especialmente en el contexto de la determinación del destino celular durante las etapas iniciales del desarrollo. Además, resulta imperativo profundizar en la comprensión de los mecanismos moleculares que rigen el mantenimiento y la diferenciación de estas células, con el propósito de superar las actuales limitaciones y lograr que sus potenciales aplicaciones sean efectivas en el ámbito clínico. Estas aplicaciones abarcan desde el modelado preciso de enfermedades hasta las innovadoras posibilidades de la medicina regenerativa.

En las últimas décadas, la investigación en pluripotencia se ha focalizado primordialmente en desentrañar los mecanismos transcripcionales y epigenéticos que participan en su regulación (Atlasi y Stunnenberg, 2017; Li y Belmonte, 2017). Contrariamente, los mecanismos de control basados en modificaciones químicas del ARN han permanecido en gran medida inexplorados. En este contexto, existen amplias evidencias que respaldan el papel crucial de ciertas modificaciones epitranscriptómicas, como m^6A , en el mantenimiento del equilibrio entre los procesos de autorrenovación y diferenciación de las ESCs (Batista et al., 2014; Geula et al., 2015; Ignatova et al., 2020b; Wang et al., 2014; Xing et al., 2020). Específicamente, la eliminación de esta marca, tanto en ARNm (Batista et al., 2014; Sun et al., 2020a; Wang et al., 2014; Wen et al., 2018), como en ARNr (Ignatova et al., 2020b; Xing et al., 2020), compromete los procesos de autorrenovación y la salida de pluripotencia. Sin embargo, se dispone de poca información sobre otras modificaciones del ARN en la regulación del estado celular pluripotente, como es el caso de la marca m^5C . La presencia de esta modificación en diversos tipos de ARN, que incluyen ARNm, ARNnc, ARNr y ARNt, ha sido documentada en ESCs murinas (Amort et al., 2017; Legrand et al., 2017; Trixl et al., 2018) y humanas (Sajini et al., 2019; Selmi et al., 2021) por diferentes estudios. Teniendo esto en cuenta, en esta tesis doctoral nos hemos centrado en desentrañar las funciones, hasta ahora desconocidas, de la m^5C -metiltransferasa de ARN DNMT2 en la regulación del estado celular pluripotente.

Por un lado, hemos demostrado, mediante dos enfoques distintos (silenciamiento mediado por shRNAs y edición génica mediante CRISPR/Cas9), que la pérdida de función de DNMT2 es dispensable para el mantenimiento de la pluripotencia en condiciones estándar de cultivo, es decir, en presencia de suero y LIF (**Figuras R14-R20**). Estos datos respaldan las observaciones detalladas por estudios previos, en los cuales la eliminación de DNMT2 es compatible con el cultivo de las ESCs. En uno de estos trabajos, se logra la derivación y expansión de estas células a partir de ratones *Dnmt2* KO (Ghanbarian et al., 2016), mientras que en el otro, la eliminación de una parte homóloga al sitio putativo catalítico del gen ortólogo a *Dnmt2* en levaduras, tampoco afecta a la propagación de las ESCs en cultivo (Okano et al., 1998). Por otro lado, nuestros resultados indican que la actividad catalítica de DNMT2 sobre el ARN es esencial para la activación de programas transcripcionales implicados en la diferenciación celular temprana (**Figuras R21, R22 y R24**). De este modo, su eliminación

compromete la especificación del linaje que tiene lugar durante la salida de pluripotencia (**Figuras R26, R28 y R35**).

En contraste con los estudios iniciales que documentaron que la pérdida de función de DNMT2 no afectaba la fertilidad y viabilidad de organismos como *Arabidopsis thaliana*, *Drosophila melanogaster* y *Mus musculus* (Goll et al., 2006), nuestros resultados concuerdan con investigaciones posteriores que muestran la importancia de la función m⁵C-metiltransferasa de DNMT2 en la diferenciación de poblaciones celulares específicas, tanto en pez cebra (Rai et al., 2007), como en ratón (Ghanbarian et al., 2016; Tuorto et al., 2015; Tuorto et al., 2012). Por ejemplo, en pez cebra, la inhibición de esta enzima compromete la diferenciación celular que ocurre en el hígado, la retina y el cerebro, observándose además deficiencias en la metilación de ARNt (Rai et al., 2007). En el caso del ratón, estudios más detallados detectaron defectos en la diferenciación de células troncales progenitoras hematopoyéticas en ratones *Dnmt2* KO de 8 días hasta un año de edad (Tuorto et al., 2015), así como el desarrollo de hipertrofia cardíaca en ratones adultos (Ghanbarian et al., 2016).

Considerando nuestros resultados en base a la literatura, la ausencia de fenotipos graves en organismos KO para *Dnmt2*, a pesar de la esencialidad de su actividad catalítica durante eventos de diferenciación temprana, podría deberse a la compensación de su función por parte de otras m⁵C-metiltransferasas, como las NSUN. Una evidencia que avala esta hipótesis es que, a pesar de que la eliminación individual de las proteínas DNMT2 (Goll et al., 2006) y NSUN2 (Blanco et al., 2011) no repercute negativamente en la viabilidad de los ratones, su ausencia conjunta sí conduce a una reducción significativa en la supervivencia de estos organismos (Tuorto et al., 2012). En concordancia, otro estudio en células humanas muestra que únicamente la pérdida de función simultánea de las m⁵C-metiltransferasas DNMT2, NSUN2 y NSUN6, y no el silenciamiento individual de cada una, reduce de manera significativa los niveles de m⁵C en las fracciones de ARNt y ARN largo, lo que sugiere la existencia de mecanismos de compensación de función entre estas tres enzimas (Selmi et al., 2021).

En estudios previos, los defectos en la diferenciación celular en ausencia de DNMT2 se han asociado principalmente a la pérdida de m⁵C en ARNt (Rai et al., 2007; Tuorto et al., 2015; Tuorto et al., 2012). Concretamente, la pérdida de función de esta proteína provocó en pez cebra una reducción significativa en los niveles de metilación del ARNt-Asp (Rai et al., 2007) y una ausencia total de m⁵C38 en ARNt-Gly-GCC, ARNt-Val-ACC y ARNt-Asp-GTC en adipocitos, osteoblastos y células de la médula ósea de ratón, asociada a defectos en la diferenciación de estos tipos celulares (Tuorto et al., 2015). Estos defectos se relacionaron con errores en la traducción de proteínas específicas, enriquecidas en Asp, ya que la pérdida de m⁵C38 impacta negativamente en la fidelidad de su traducción al reducir la capacidad discriminatoria entre codones altamente similares de los ARNt cargados (Tuorto et al., 2015). En línea con estas evidencias, nuestros resultados muestran la ausencia de cambios globales en la metilación de ARN total y ARN largo tras la pérdida de función de DNMT2 (**Figura R29**). Esto concuerda con el análisis de los datos de RNA BS-seq de ESCs *Dnmt2* WT y KO previamente publicados (Legrand et al., 2017), donde no se observan diferencias en los niveles de m⁵C en las fracciones de ARNr y ARN largo, pero sí en ARNt (**Figura R30d**). Considerando esto y que la expresión de los transcritos largos marcados con m⁵C no varía tras la eliminación de DNMT2 (**Figura R30c**), los resultados obtenidos sugieren que las alteraciones transcripcionales observadas en ESCs *Dnmt2* KO y durante su diferenciación pueden estar directa o indirectamente relacionadas con la pérdida de m⁵C38 en moléculas de ARNt. De este modo, la deficiencia observada en la expresión de marcadores de diferenciación durante la especificación temprana de ESCs *Dnmt2*

KO sugiere la existencia de un mecanismo de regulación transcripcional mediado por la metilación de ARNt dependiente de DNMT2.

En base a las evidencias previamente mencionadas, proponemos un potencial mecanismo de regulación transcripcional en ESCs mediado por la actividad catalítica de DNMT2 sobre ARNt (**Figura D2**). La pérdida de esta actividad se asocia con un aumento en la generación de 5'-tsRNAs derivados de ARNt-Gly-GCC, ARNt-Val-ACC y ARNt-Asp-GTC, tanto en células murinas de la médula ósea (Tuorto et al., 2015), como espermáticas (Yu et al., 2021; Zhang et al., 2018), así como en *Drosophila melanogaster* (Durdevic et al., 2013b; Schaefer et al., 2010). Estas moléculas son capaces de regular la expresión génica mediante la interacción con proteínas de unión a ARN, como se ha observado en células humanas tumorales (Goodarzi et al., 2015; Maute et al., 2013), no tumorales (Kuscu et al., 2018) y en ESCs murinas (Krishna et al., 2019). De hecho, según un estudio publicado en 2019, los tsRNAs parecen desempeñar un papel crítico en la modulación del estado celular pluripotente (Krishna et al., 2019). Los autores demuestran que la inhibición de estas moléculas, mediante el uso de oligonucleótidos antisentido, resulta en la regulación al alza de algunos marcadores de pluripotencia, como *Klf4*, en condiciones de diferenciación (Krishna et al., 2019). En contraposición, observan que la interacción entre un 5'-tsRNA (5'-tsGlnCTG) y la proteína de unión al ARN IGF2BP disminuye los niveles del transcrito *c-Myc* durante la diferenciación de las ESCs. De manera destacada, en este mismo trabajo también se muestra un aumento significativo de los niveles de 5'-ts-Gly-GCC, tsRNA asociado a la pérdida de función de DNMT2, tras 6 días en condiciones de diferenciación. En consonancia, nosotros observamos una reducción significativa en la expresión de *Dnmt2* en el día 6 del proceso de generación de EBs (**Figura R26b**). Conjuntamente, estos datos podrían sugerir que, durante un correcto proceso de diferenciación, la función de DNMT2 es necesaria en etapas tempranas del proceso de salida de pluripotencia. Sin embargo, una vez establecida la especificación del linaje (día 6 del proceso), el descenso en los niveles de DNMT2 provocaría un incremento en la generación de tsRNAs, tal y como documentan Krishna y colaboradores. De este modo, deberíamos evaluar si una acumulación prematura de tsRNAs en ESCs deficientes en DNMT2 podría conducir a una inhibición aberrante de transcritos necesarios para el inicio de este proceso (**Figura D2**). Para confirmar esta hipótesis, podríamos, por un lado, validar por *northern blot* la generación de 5'-tsRNAs derivados de ARNt-Asp, ARNt-Gly y ARNt-Val durante la diferenciación de ESCs KO para *Dnmt2*. En este caso, esperaríamos un incremento aberrante en la generación de estas moléculas en las etapas iniciales de la diferenciación de células deficientes en DNMT2, en comparación con ESCs *Dnmt2* WT. Por otro lado, deberíamos validar que la inhibición de los 5'-tsRNAs generados consigue restaurar los niveles adecuados de expresión de los marcadores tempranos de diferenciación durante la salida de pluripotencia de las ESCs *Dnmt2* KO. Además, también deberíamos confirmar que la expresión de la forma WT y no la catalíticamente inactiva de DNMT2 (C79A) en células *Dnmt2* KO consigue restaurar un nivel adecuado de los tsRNAs analizados durante etapas tempranas del proceso de diferenciación.

A pesar de las alteraciones transcripcionales observadas en ausencia de DNMT2, no podemos descartar que la pérdida m⁵C38 asociada (**Figura R30**) repercuta negativamente en la traducción de proteínas reguladoras del estado celular pluripotente. Dada la regulación dinámica de las modificaciones del ARNt con el fin de ajustar la síntesis de proteínas a las necesidades celulares (Gingold et al., 2014), y que las células troncales, ya sean embrionarias (Bornelöv et al., 2019; Sampath et al., 2008) o adultas (Blanco et al., 2016; Signer et al., 2014), son especialmente sensibles a cambios en el proceso de traducción, cabe esperar que la

desregulación de m⁵C-metiltransferasas de ARNt resulte en alteraciones del estado celular pluripotente.

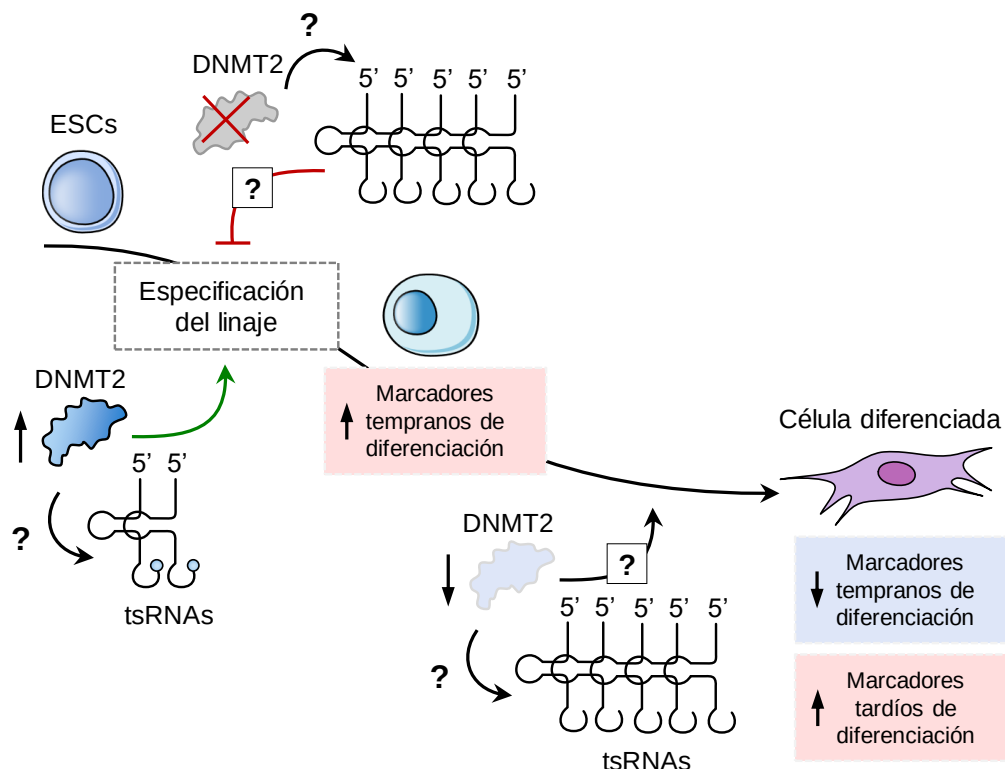


Figura D2. Potencial mecanismo según el que DNMT2 promueve la especificación del linaje mediante la metilación de ARNt. En esta imagen se ilustra la hipótesis planteada respecto a la función catalítica dependiente de DNMT2 durante la salida de pluripotencia. Según nuestros resultados, los niveles de expresión de *Dnmt2* se mantienen elevados en las etapas iniciales de la salida de pluripotencia y descienden una vez establecida la especificación del linaje. Además, mostramos que la actividad catalítica en ARN de esta proteína es necesaria para la activación de genes implicados en la diferenciación temprana, de modo que su ausencia compromete la adecuada especificación de linaje tras la salida del estado celular pluripotente. Basándonos en la literatura relacionada con otros contextos celulares (Schaefer et al., 2010; Tuorto et al., 2015; Zhang et al., 2018), la pérdida de la función catalítica de DNMT2 resulta en la pérdida de m⁵C38 y en una acumulación aberrante de 5'-tsRNAs. Durante la diferenciación temprana, esto podría comprometer la activación de los programas transcripcionales involucrados. Además, otro estudio llevado a cabo en ESCs (Krishna et al., 2019) muestra la acumulación 5'-tsRNAs en etapas más tardías del proceso de diferenciación, cuando la expresión de los marcadores de especificación del linaje ya se ha producido, lo que concuerda con el descenso observado en los niveles de *Dnmt2*. De ser así, los fragmentos generados podrían participar en la inhibición de marcadores tempranos de diferenciación, favoreciendo así la expresión de marcadores tardíos. El color verde indica el favorecimiento del proceso y el color rojo implica la inhibición de este. Los signos de interrogación indican incógnitas no resueltas. Esta imagen fue elaborada mediante el uso del programa Inkscape (v1.3).

En consonancia con nuestras observaciones, la actividad catalítica de las m⁵C-metiltransferasas de ARNt, NSUN3 (Trixl et al., 2018) y NSUN2 (Blanco et al., 2011; Hussain et al., 2013b), resulta crítica para el mantenimiento del equilibrio entre los procesos de autorrenovación y diferenciación en células troncales.

De manera similar a DNMT2, la pérdida de la actividad catalítica de NSUN3 no afecta al mantenimiento de ESCs murinas en cultivo, aunque sí resulta en una reducción significativa de su proliferación (Trixl et al., 2018). Mientras que la eliminación de DNMT2 conduce a una reducción en la activación de los marcadores de las tres capas germinales durante la formación de EBs (Figura R26), la pérdida de actividad catalítica de NSUN3 provocaba un sesgo de diferenciación hacia los linajes mesodermo y endodermo en detrimento del neuroectodermo

(Trixl et al., 2018). Además, no solo se observan cambios a nivel transcripcional, sino también en la síntesis de proteínas mitocondriales, asociados a la reducción en la metilación del ARNt mitocondrial que porta metionina (Trixl et al., 2018).

En concordancia, la metilación de ARNt por parte de NSUN2 es necesaria para la diferenciación de las células troncales adultas en la epidermis (Blanco et al., 2016; Blanco et al., 2011), resultando su pérdida en la acumulación de 5'-tsRNAs que alteran el proceso de traducción de estas células (Blanco et al., 2016).

Respecto a NSUN6, de manera similar a nuestros datos obtenidos para DNMT2, aunque su eliminación no altera la pluripotencia de ESCs humanas, sí altera la expresión de algunos marcadores de diferenciación temprana durante la diferenciación a EBs, como *Gata6* o *Cdx2* (Selmi et al., 2021).

Conjuntamente, la actividad m⁵C-metiltransferasa sobre el ARNt parece jugar un papel relevante en la determinación del destino celular en células troncales embrionarias y adultas, tanto a nivel transcripcional como de síntesis de proteínas. Por lo tanto, aunque observamos cambios a nivel transcripcional en ESCs KO para *Dnmt2*, no podemos descartar que la pérdida de m⁵C38 en estas células también resulte en la alteración de la síntesis de proteínas implicadas en el proceso de diferenciación celular.

Dado el poder de resolución de las técnicas mediante las que mostramos la ausencia de cambios globales en la metilación de otros ARNs en ESCs *Dnmt2* KO (**Figura R29 y R30**), tampoco podemos descartar su actuación específica sobre otras especies de ARN. De hecho, en ESCs murinas se ha descrito que esta proteína cataliza la deposición de m⁵C en el ARNnc *Rn7SK*, que inhibe la activación de la ARN polimerasa II al asociarse con el factor regulador de la elongación de la transcripción, P-TEFb. En consonancia con nuestros datos, este estudio muestra que DNMT2 es importante durante la salida de pluripotencia, aunque en este caso, asocian su función a la metilación de ARNnc y no de ARNt (Ghanbarian et al., 2016).

En última instancia, es relevante destacar que nuestros hallazgos indican que DNMT2 desempeña un papel relevante en la regulación transcripcional de subfamilias específicas de retrotransposones en ESCs (**Figura R25**). Estos resultados respaldan investigaciones previas que han asociado la pérdida de función de DNMT2 con la desrepresión de elementos transponibles (TEs) en *Drosophila melanogaster* (Genencher et al., 2018; Phalke et al., 2009). En esta especie, se ha demostrado que la expresión de una forma de DNMT2 catalíticamente inactiva es suficiente para restablecer los niveles normales de estos elementos (Genencher et al., 2018). A mayores, otras modificaciones químicas del ARN, como m⁶A, han sido extensamente vinculadas al control de retrotransposones en ESCs (Chelmicki et al., 2021; Chen et al., 2021a; Liu et al., 2020a; Liu et al., 2021). Aunque este control es esencial para mantener la identidad y la capacidad de diferenciación de las ESCs (Fu et al., 2021; Hackett et al., 2017), aún no sabemos si DNMT2 participa en la regulación del estado celular pluripotente mediante la metilación de este tipo de elementos. Para verificar si la función de DNMT2 en la regulación transcripcional específica de subfamilias de retrotransposones depende de su actividad catalítica, aprovecharemos los datos de secuenciación masiva generados en esta tesis doctoral, que incluyen las ESCs WT y KO para *Dnmt2*, así como las ESCs KO rescatadas con las formas mutantes inactivas (C79A y R371H) (**Figura R33**). Realizaremos un análisis detallado para determinar la dependencia o independencia de la actividad catalítica de DNMT2 en la modulación de los TEs. Posteriormente, validaremos los cambios identificados en los niveles de expresión de estos elementos mediante RT-qPCR. Este enfoque nos permitirá comprender

mejor el papel específico de la actividad catalítica de DNMT2 en la regulación de retrotransposones en ESCs.

7.3 EL PAPEL DE DNMT2 EN CÁNCER DE MAMA

La alteración de los patrones de modificación química del ARN se ha asociado con el desarrollo tumoral y con procesos que incluyen la proliferación, la migración y la capacidad de invasión celular, habitualmente desregulados en cáncer (Nombela et al., 2021). Aunque las mutaciones en metiltransferasas de ARN no suelen ser consideradas como inductoras directas de cáncer, la función de muchas de estas proteínas se ha relacionado funcionalmente con la supervivencia, el crecimiento y la diferenciación de las células tumorales (Chen et al., 2019; He et al., 2020; Ignatova et al., 2020a; Mei et al., 2020; Sun et al., 2020b). De acuerdo con investigaciones anteriores (Du et al., 2020; Elhardt et al., 2015), hemos caracterizado que el gen *DNMT2* se encuentra mutado en una amplia variedad de tipos de cáncer (**Figura R39**). Esto contrasta con el hecho de que pocos estudios hayan profundizado en su papel en un contexto celular tumoral (Adamczyk-Grochala et al., 2023; Betlej et al., 2022; Bloniarz et al., 2021; Sun et al., 2021). Dado que nuestros resultados sugieren que DNMT2 participa en la regulación del fenotipo de células de cáncer de mama mediante su actividad catalítica, los datos aquí obtenidos contribuyen a definir su papel en la biología de este tipo de cáncer.

En este trabajo, demostramos que la pérdida de función de DNMT2 generalmente se correlaciona con un peor pronóstico en pacientes con diversos tipos de cáncer. Tanto las mutaciones en el gen *DNMT2* como la disminución de sus niveles de expresión se asocian con una reducción en la supervivencia de los pacientes (**Figura R41**). En concordancia con estos datos, análisis previos han sugerido que DNMT2 actúa como un factor protector en la supervivencia de pacientes con algunos tipos de cáncer, como el adenocarcinoma pulmonar (Liu et al., 2022) o el cáncer de recto (Zhang et al., 2023). Aquí, mostramos que las mutaciones en *DNMT2* se acumulan frecuentemente en tumores mamarios, produciéndose una reducción significativa de su expresión en el tumor respecto al tejido normal (**Figura R42**). En este contexto, una menor expresión de *DNMT2* también correlaciona con un peor pronóstico en pacientes con cáncer de mama, especialmente en subtipos poco agresivos (luminales A y B), sin observarse diferencias significativas en pacientes con subtipos HER2+ o triple negativo (**Figura R43**).

Los análisis realizados en líneas celulares tumorales respaldan estos resultados, ya que la pérdida de función de DNMT2 provoca un aumento en la agresividad de la línea celular MCF-7 (subtipo luminal A) pero no de la línea MDA-MB-231 (subtipo triple negativo) (**Figuras R44 y R45**). En MCF-7, la expresión de *DNMT2* muestra una correlación negativa con la de *MDM2*, el principal regulador negativo de P53 (**Figura R46**) (Levine, 2020). Además, se observó una reducción significativa de los niveles de la proteína P53 en estas células después de la inhibición de *DNMT2* (**Figura R46**). Este hallazgo podría explicarse por el incremento en la expresión de *MDM2* tras el silenciamiento de *DNMT2* (**Figura R46**), lo que resultaría en un aumento de la presencia de *MDM2* y, por ende, en una disminución de los niveles proteicos de P53 (**Figura R46**). Además, también mostramos que la expresión de *MDM2* disminuyó significativamente en células MCF-7 después de la sobreexpresión de la forma catalíticamente activa de DNMT2, pero no de la mutante inactiva (C79A) (**Figura R46**). Considerando estos resultados, sugerimos que DNMT2 ejerce una regulación positiva sobre los niveles de P53 en células MCF-7 y participa en la regulación transcripcional de *MDM2* a través de su actividad catalítica.

Sin embargo, aún desconocemos el mecanismo molecular a través del cual la modificación m^5C mediada por DNMT2 puede participar en la regulación del eje MDM2-P53. Esta marca epitranscriptómica ha sido asociada con la estabilización de transcritos en diversos tipos de cáncer (Chen et al., 2021b; Chen et al., 2019; Yu et al., 2023). Por ejemplo, en cáncer de vejiga, esta estabilización de transcritos modificados es promovida por el "reader" de m^5C , YBX1, cuya inhibición en cáncer de mama se ha asociado con una menor agresividad de las células tumorales (Goodarzi et al., 2015; Lim et al., 2017; Lin et al., 2022). Teniendo esto en cuenta, no esperaríamos una metilación directa del transcrito *MDM2* por parte de DNMT2, ya que, basándonos en la literatura donde esta modificación ha sido ampliamente asociada con la estabilización de transcritos en múltiples contextos celulares (Chen et al., 2021b; Chen et al., 2019; Liu et al., 2022a; Yang et al., 2019c; Yu et al., 2023), su silenciamiento conduciría a una desestabilización de este ARN y no a una regulación al alza del mismo. Es por esta razón, que planteamos la posibilidad de que la regulación transcripcional de *MDM2* mediada por DNMT2 tenga lugar indirectamente a través de la metilación de ARNt (**Figura D3**).

La generación de tsRNAs ha sido previamente relacionada con la regulación de la tumorigénesis en distintos tipos de cáncer (Balatti et al., 2017), entre los que se incluye el cáncer de mama (Dhahbi et al., 2014; Goodarzi et al., 2015; Honda et al., 2015; Wang et al., 2020a). Concretamente, en células MDA-MB-231 se ha observado que estas moléculas pueden modular la expresión de transcritos mediante la interacción con la proteína YBX1 (Goodarzi et al., 2015). No obstante, los patrones de expresión de los tsRNAs varían en distintos tipos de tumor, sugiriendo que en función del contexto estas moléculas pueden actuar como promotoras o como supresoras de tumores (Balatti et al., 2017; Honda et al., 2015).

Teniendo esto en cuenta, junto con la reducción en los niveles de *DNMT2* en tumores mamarios respecto al tejido normal (**Figura R39**), cabe la posibilidad de que el agravamiento del fenotipo observado en células MCF-7 se deba a la acumulación de tsRNAs y que los cambios transcripcionales observados en *MDM2* se encuentren también mediados por la generación de estos fragmentos (**Figura D3**). Como ya se ha visto en células de cáncer de mama, los tsRNAs pueden modular la expresión génica mediante la interacción con proteínas de unión a ARN (Goodarzi et al., 2015), aunque en este caso esperaríamos la interacción de proteínas que actúen como reguladores negativos de *MDM2*. Para confirmar que la regulación al alza de *MDM2* observada tiene lugar a través de la generación de 5'-tsRNAs, podríamos inhibir estas moléculas mediante el empleo de oligonucleótidos antisentido en células MCF-7 deficientes en DNMT2. De este modo, no esperaríamos el aumento previamente observado en la expresión de este gen.

En consonancia con la hipótesis planteada, el incremento en los niveles de estas moléculas se ha asociado con una mayor proliferación en células de cáncer de mama (Honda et al., 2015). De hecho, la generación de algunos de estos fragmentos resulta en el favorecimiento de la colonización pulmonar metastásica del cáncer de mama (Honda et al., 2015). Curiosamente, tanto en cánceres luminales como en células MCF-7 se ha detectado un aumento en la generación de 5'-tsRNAs derivados de ARNt-Asp-GTC, uno de los principales sustratos de DNMT2 (Honda et al., 2015). Si la inhibición de *DNMT2* condujese a la generación de 5'-tsRNAs en células MCF-7, los niveles de estos fragmentos observados en estas células (Honda et al., 2015) podrían verse incrementados y provocar un agravamiento del fenotipo, tal y como observamos en este trabajo (**Figuras R43-R45**).

Por otro lado, se ha documentado en varias líneas celulares tumorales, que incluyen células MCF-7, que mayores niveles de MDM2 pueden actuar promoviendo la transición

epitelial mesenquimal (EMT) (Ou et al., 2021; Tang et al., 2019), un proceso inverso a la MET, con importantes implicaciones en la progresión tumoral, la metástasis y la resistencia a la quimioterapia (Lu y Kang, 2019). Teniendo esto en cuenta y que nuestros resultados durante la SCR muestran que la pérdida de función de DNMT2 promueve el mantenimiento de características mesenquimales, cabe la posibilidad de que el silenciamiento de *DNMT2* favorezca la EMT mediante la regulación al alza de *MDM2*. Aunque desconocemos si las funciones de DNMT2 observadas en el contexto de reprogramación celular se encuentran conservadas en células tumorales, el elevado grado de conservación del papel de DNMT2 en distintos tipos celulares (por ejemplo, células de la médula ósea, espermáticas y adipocitos, entre otras) y organismos (*Drosophila melanogaster*, *Mus musculus* y *Homo sapiens*) sugiere que esta función podría estar conservada.

Es preciso destacar que de manera opuesta a lo observado en este trabajo y en otros (Betlej et al., 2022; Bloniarz et al., 2021), algún estudio documenta una correlación positiva entre los niveles de DNMT2 y el desarrollo tumoral (Sun et al., 2021). Es más, su eliminación en algunos tipos de cáncer sensibiliza a las células a terapias antitumorales (Adamczyk-Grochala et al., 2023; Chen et al., 2020). Esto podría indicar que el papel de DNMT2 depende del tejido y del contexto, o que una regulación precisa de sus niveles es crítica en estas células para evitar la progresión del tumor. No obstante, la totalidad de las funciones de la modificación m^5C dependiente de DNMT2 en el cáncer aún está por descubrir.

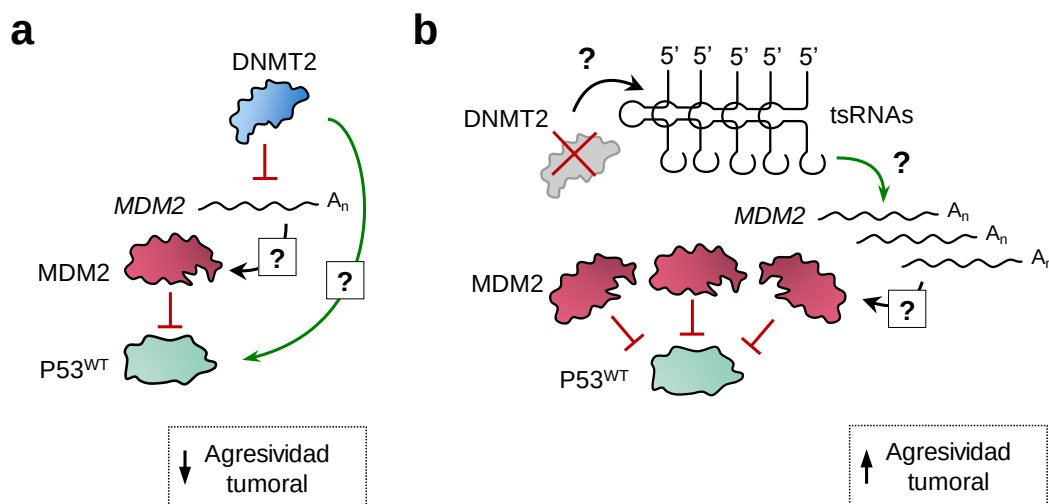


Figura D3. Potencial mecanismo mediante el cual el silenciamiento de *DNMT2* conlleva un fenotipo más agresivo de la línea tumoral MCF-7 de cáncer de mama. Esta ilustración representa la hipótesis planteada según la cual DNMT2 regula el fenotipo de células MCF-7 mediante la metilación de ARNt. **a)** En base a nuestros resultados, mayores niveles de DNMT2 correlacionan con una menor agresividad tumoral de la línea celular MCF-7. En estas células, DNMT2 regula negativamente la expresión de *MDM2* y positivamente los niveles de la proteína P53. Aunque desconocemos si esta regulación ocurre de manera directa o indirecta. **b)** Por el contrario, la pérdida de función de DNMT2 agrava el fenotipo tumoral de estas células. En este contexto, DNMT2 parece regular la expresión de *MDM2* mediante su actividad catalítica, aunque desconocemos el sustrato sobre el cual esta proteína cataliza la deposición de m^5C . Basándonos en evidencias previas en el contexto del cáncer de mama (Goodarzi et al., 2015; Honda et al., 2015), que muestran un papel relevante de la generación de tsRNAs en este proceso, planteamos la hipótesis de que DNMT2 regula el fenotipo de las células MCF-7 a través de la metilación de ARNt. De este modo, el silenciamiento de *DNMT2* podría conducir a un aumento en la generación de tsRNAs, con capacidad de modular la expresión génica y cuyo impacto en la regulación transcripcional de *MDM2* desconocemos por el momento. Una mayor transcripción y traducción de *MDM2* conduciría a una mayor degradación de la proteína P53, lo que resultaría en la disminución de sus niveles totales de proteína. El color verde indica el favorecimiento del proceso y el rojo su inhibición. Las incógnitas planteadas se indican con un signo de interrogación. Esta imagen fue elaborada mediante el uso del programa Inkscape (v1.3).

8. CONCLUSIONS / CONCLUSIONES

1. DNMT2 is necessary for the formation of iPSCs through OSKM in mesenchymal somatic cells.

DNMT2 es necesario para la formación de iPSCs a través de OSKM en células somáticas mesenquimales.

2. The m⁵C-methyltransferase activity of DNMT2 on RNA is essential for the acquisition of the epithelial cell fate during MET.

La actividad m⁵C-metiltransferasa de DNMT2 en el ARN es esencial para la adquisición del destino celular a epitelial durante la MET.

3. The requirement for DNMT2 function in direct reprogramming to iHSCs is conserved across species.

La necesidad de la función de DNMT2 en la reprogramación directa a iHSCs está conservada entre especies.

4. DNMT2 absence is compatible with pluripotency maintenance.

La ausencia de DNMT2 es compatible con el mantenimiento de pluripotencia.

5. DNMT2 is involved in the regulation of transcriptional programs in ESCs, both dependently and independently of its catalytic activity on RNA.

DNMT2 participa en la regulación de programas transcripcionales en ESCs, tanto de manera dependiente como independiente de su actividad catalítica en el ARN.

6. The catalytic activity of DNMT2 on RNA is necessary for a proper exit from the pluripotent cellular state.

La actividad catalítica de DNMT2 en el ARN es necesaria para una salida adecuada del estado celular pluripotente.

7. Mutations in *DNMT2* or the reduction of its expression correlate with an unfavorable prognosis in patients affected by various types of cancer, including breast cancer.

Las mutaciones en *DNMT2* o la disminución de su expresión se correlacionan con un pronóstico desfavorable en pacientes afectados por diversos tipos de cáncer, incluyendo el cáncer de mama.

8. The silencing of *DNMT2* intensifies the aggressiveness of the MCF-7 tumor cell line, belonging to the luminal A subtype of breast cancer.

El silenciamiento de *DNMT2* intensifica la agresividad de la línea tumoral MCF-7, perteneciente al subtipo luminal A de cáncer de mama.

9. APÉNDICE

9.1. Reactivos, kits y material fungible

Tabla suplementaria 1. Reactivos y químicos.

Nombre	Marca	Referencia
3xFLAG-dCas9/pMXs-neo	Addgene	51260
4',6-diamidino-2-phenylindole (DAPI)	Sigma-Aldrich	D9542
Acetona	Fisher Chemical	A/0560/17
Ácido acético glacial	Fisher Scientific	10171460
Ácido retinoico (RA)	Santa Cruz Biotechnology	sc-200898
Acrilamida/bisacrilamida	Sigma-Aldrich	A3574
Adenosina	Sigma-Aldrich	A4036
Agar	Sigma-Aldrich	A5431
Agua libre de ribonucleasas	Invitrogen	10977-035
Albúmina de suero bovino (BSA)	Fisher Bioreagents	BP1605
Aminoácidos no esenciales	Gibco	11140-050
Ampicilina	Fisher Bioreagents	BP1760
ATP pH7	TCI Chemicals	A0157
B-27	Gibco	17504044
β-mercaptoetanol	Sigma-Aldrich	M6250
BMP4	Gibco	PHC9534
Bolas magnéticas (Dynabeads Protein G)	Invitrogen	1004D
Azul de bromofenol	Sigma-Aldrich	B8026
CaCl ₂ ·2H ₂ O	Fisher Chemical	VL2151000
Cebadores hexaméricos aleatorios	Thermo Scientific	SO142
CHIR99021	Sigma-Aldrich	SML1046
Citidina	Sigma-Aldrich	C4654
Cloroformo	Fisher Chemical	C/4920/15
Cocktail inhibidor de proteasas (PIC)	Sigma-Aldrich	P8340
Cristal violeta	Sigma-Aldrich	C0775-25G
DTT	Sigma-Aldrich	43816
EDTA	Fisher BioReagents	BP2482
EGTA	Alfa Aesar	J61721
Enzima de restricción <i>AgeI</i>	Thermo Scientific	ER1461
Enzima de restricción <i>BbsI</i>	Thermo scientific	ER1012
Enzima de restricción <i>BsaI</i>	Thermo scientific	ER0291
Enzima de restricción <i>Clal</i>	Thermo Scientific	FD0144
Enzima de restricción <i>EcoRI</i>	Thermo Scientific	ER0271
Enzima de restricción <i>NheI</i>	Thermo Scientific	ER0971
Enzima de restricción <i>NotI</i>	Thermo Scientific	FD0594
Enzima de restricción <i>PacI</i>	Thermo Scientific	ER2202
Etanol	Sigma-Aldrich	1.08543
Factor Inhibidor de la Leucemia (LIF)	producido en el laboratorio	n/a
Fenol:cloroformo:isoamil alcohol (PCIA 25:24:1)	Affymetrix	75831
Formaldehido 37%	Sigma-Aldrich	F8775
Fosfatasa alcalina termosensible FasAP	Thermo Scientific	EF0654
G418	Corning	30-234-CR
Gelatina	Sigma-Aldrich	G9391
Glicerol	Sigma-Aldrich	G5516
Glicina	Sigma-Aldrich	G8898
Glucógeno	Thermo Scientific	R0561
Guanosina	Sigma-Aldrich	G6264
HEPES pH 7,9	Fisher Bioreagents	BP299
Hidrocortisona	STEMCELL Technologies	74142
Higromicina	Corning	30-240-CR
Igepal	Sigma-Aldrich	I3021
Isopropanol	Fisher Bioreagents	BP2618
Kanamicina	Fisher Bioreagents	BP906-5
KCl	Fisher Chemical	P/4240/60
Leche en polvo	Lab Scientific	M0841
L-Glutamina	Gibco	25030-081
Ligasa de ADN T4	Thermo Scientific	EL0016

Lipofectamina 3000	Invitrogen	L3000-015
Medio Eagle modificado de Dulbecco (DMEM) alto en glucosa	Corning	10-013-CV
Medio Eagle modificado de Dulbecco (DMEM) alto en glucosa (polvo)	Gibco	52100-021
Medio DMEM con mezcla de nutrientes F-12 Ham (DMEM/F-12)	Sigma-Aldrich	D6421
Medio Neurobasal	Gibco	21103049
Mezcla de dNTPs	Thermo Scientific	R0191
Mezcla maestra de polimerasa de ADN DreamTaq Green PCR	Thermo Scientific	K1081
Mezcla maestra de polimerasa de alta fidelidad Phusion Green Hot Start II	Thermo Scientific	F5665
Mezcla maestra de polimerasa de alta fidelidad Platinum SuperFi II PCR	Invitrogen	12368010
MgCl ₂	Sigma-Aldrich	M1028
MnCl ₂ ·4H ₂ O	Sigma-Aldrich	M8266
Myelocult H5100	STEMCELL Technologies	05150
Myelocult M5300	STEMCELL Technologies	05350
N-2	Gibco	17502048
NaCl	Fisher Bioreagents	BP358-1
NaFl	Acros Organics	201290250
NaHCO ₃	Thermo Scientific	10128190
Ortovanadato	MP Biomedicals	215966410
Paraformaldehído	Sigma-Aldrich	P6148
DPBS con Ca ⁺² y Mg ⁺²	Sigma-Aldrich	D8662
PD0325901	Sigma-Aldrich	PZ0162
Penicilina/estreptomicina	Gibco	15070-063
pFETCh-3xFlag-P2A-NeoR	Addgene	63934
PIPES pH 6.7	Sigma-Aldrich	P6757
pLKO.1 puro	Addgene	8453
pMDG2	Addgene	12259
PMSF	Santa Cruz Biotechnology	sc-482875
Polibreno	Sigma-Aldrich	H9268
Polietilenimina (PEI)	Sigma-Aldrich	408727
PowerUp SYBR Green	Applied Biosystems	A25741
pPB-CAG-3xFLAG-pA-pgk-hph	Addgene	48753
Proteinasa K	Thermo Scientific	EO0491
psPAX2	Addgene	12260
Puromicina	Fisher Bioreagents	BP2956
qScript cDNA SuperMix	Quantabio	P/N 84034
Reactivo de Ponceau S	Sigma-Aldrich	P3504
Reactivo TRIzol	Invitrogen	15596018
Ribonucleasa A	Thermo Scientific	EN0531
SDS	Fisher Bioreagents	BP1311
SpCas9-2A-Puro V2.0	Addgene	62988
Suero fetal bovino (FBS)	Gibco	10270106
SuperScript IV RT	Invitrogen	18090010
Sustrato para inmunotransferencia ECL Pierce	Thermo Scientific	32106
T4 PNK	Thermo Scientific	EK0032
Tampón de reacción A de la T4 PNK	Thermo Scientific	EK0032
Tampón de reacción NEB	New England BioLabs	B70025
Tampón fosfato salino de Dulbecco (DPBS)	Corning	21-031-CV
Tampón salino Tris	Fisher Bioreagent	BP2471-1
TFG-8	Sigma-Aldrich	H8541
Timidina	Sigma-Aldrich	T1895
Transcriptasa inversa Maxima H Minus	Thermo Scientific	EP0753
Tripsina-EDTA	Sigma-Aldrich	T3924
Tris-HCl pH 6,8	Bio Basic	SD8122
Tris-HCl pH 7,5	Fisher Bioreagents	BP1757
Triton X-100	Sigma-Aldrich	T8787
TWEEN 20	Sigma-Aldrich	P1379
Uridina	Sigma-Aldrich	U3003
Yoduro de propidio	Molecular Probes	P3566

Tabla suplementaria 2. Kits.

Nombre	Marca	Referencia
Centra Puregene Cell	Qiagen	158043
CloneJet PCR Cloning	Thermo Scientific	K1231
E.Z.N.A Total RNA kit 1	Omega Bio-Tek	R6834-02
EZ RNA Methylation	Zymo Research	R5001
Ensayo de proteínas BCA Pierce	Thermo Scientific	23227
Fraccionamiento subcelular de proteínas	Thermo Scientific	78840
GeneArt™ Gibson Assembly HiFi	Invitrogen	A46627
GenElute Gel Extraction	Sigma-Aldrich	NA1111
GenElute PCR Clean-up	Sigma-Aldrich	NA1020
GenElute Plasmid Miniprep	Sigma-Aldrich	PLN350
jetPRIME	Polyplus	114-15
Leucocyte Alkaline Phosphatase	Sigma-Aldrich	86R-1KT
MethylFlash 5-mC RNA Methylation ELISA Easy Kit	Epigentek	P-9009-48
Nucleoside Digestion Mix	New England BioLabs	M0649S
PureLink HiPure Plasmid Filter Midiprep	Invitrogen	K210015
Qubit HS Assay	Invitrogen	Q32855
Qubit RNA IQ Assay	Invitrogen	Q33222
RiboMinus Eukaryote Probe Mix v2	Ambion	A15020
RiboZero	Illumina	20040526
RNA clean and concentrator-5	Zymo Research	R1015
Tampón de lisis RIPA	Santa Cruz Biotechnology	sc-24948
Turbo DNA-free kit	Invitrogen	AM1907

Tabla suplementaria 3. Accesorios y fungible.

Nombre	Marca	Referencia
0,22 µm EMD Millipore Steriflip	Millipore	SCGP00525
Agitador de balanceo ProBot Rcoker 25	Labnet	S2025-B
Agitador de incubación	Infors HT	Minitron
Agujas de 18G ½	Terumo	8AN1838S1
Agujas de 21G ½	Terumo	8AN2138R1
Baño térmico serie VWB2	VWR	462-0556
Cámara de Neubauer	Hirschmann	8100103
Casetes de diálisis de Slide-A-Lyzer G2, 10K MWCO	Thermo Scientific	87731
Centrífuga refrigerada	MPW	260R
Centrífuga Sorvall	Thermo Scientific	ST 16R
Citómetro	BD Biosciences	FACSCalibur
Espectrofotómetro Nanodrop 2000	Thermo Scientific	ND-2000
Filtros Amicon Ultra de 0,5 mL	Sigma-Aldrich	UFC500308
Filtros Amicon Ultra-15	Millipore	UFC903024
Filtros de 0,22 µm	Fisherbrand	15181489
Fluorómetro Qubit	Invitrogen	Q33226
Imán DynaMag-2	Invitrogen	12321D
Incubador de CO2 New Brunswick	Eppendorf	Galaxy 170S
Incubadora digital INCU-Line	VWR	390-0384
Jeringuilla	Terumo	0197
Lector de placas Mithras	Berthol Technologies	LB 940
Membranas de fluoruro de polivinilideno	Cytiva	10600023
Mezclador vórtex	Fisherbrand	11746744
Microscopio invertido	Leica	DMI6000 B
Microscopio invertido con una cámara Olympus DP72	Olympus	IX51
Mini Gel Tank	Invitrogen	A25977
Novex WedgeWell 4-20% tris-Glycine Gel	Invitrogen	XP04205BOX
Novex WedgeWell 4-20% tris-Glycine Gel (10 pocillos)	Invitrogen	XP04200BOX
Películas de rayos X Fuji	Fujifilm	100NIF
Películas PlateMax UltraClear	Axygen	UC-500
Placas 150 mm de diámetro	Corning	430599
Placas de 100 mm de diámetro	Corning	430167
Placa de 6 pocillos	Corning	3516
Placa de 12 pocillos	Corning	3513

Placa de 24 pocillos	Corning	3524
Placa de 48 pocillos	Corning	3548
Placas de 96 pocillos con fondo en U	Corning	3788
Placas de reacción de 384 pocillos	Axygen	PCR-384M2-C
Placas Petri de baja adherencia	Deltalab	200200
Procesador de películas Fuji	Fujifilm	FPM-100A
Qubit assay tubes	Invitrogen	Q32856
Rotador	Grant-bio	PTR-35
Sistema de electroforesis horizontal kuroGEL	VWR	700-0034
Sistema de electroforesis vertical Mini-PROTEAN Cell	Bio-Rad	1658004
Sistema de filtrado de 1 L	Corning	431098
Sistema de filtrado de 250 mL	Corning	431096
Sistema de filtrado de 500 mL	Corning	431097
Sistema para qPCR QuantStudio 5	Applied Biosystems	A33621
Sonicador digital	Branson	Sonifier
Termobloque	Thermo Scientific	88870008
Termociclador SimpliAmp	Applied Biosystems	A24811
Transiluminador	Bio-Rad	Gel Doc XR+ Molecular Imager
Tubos 5PRIME Phase Lock Gel Heavy	Quantabio	2302830
Tubos de citometría	Corning	352058
Tubos DNA LoBind	Eppendorf	22431021
Tubos Phasemaker	Invitrogen	A33248
Tubos Protein LoBind	Eppendorf	320108132

9.2 Secuencias de oligonucleótidos

Tabla suplementaria 4. Guías de ARN para uso del sistema CRISPR/Cas9

Dnmt2 KO		
Gen	Secuencia directa 5'→3'	Secuencia inversa 5'→3'
Dnmt2	CACCGAGCTGTACAGTGGCATTGG	AAACCAATGCCACTGTACAGCTC
Dnmt2 KI		
Dnmt2	CACCGTCCCTTACATTTATGTAC	AAACGTGACATAAATGTAAGGGAC

Tabla suplementaria 5. Cebadores para la generación de shRNAs.

Gen	Secuencia
mDnmt2#1	(D) CCGGCAATGCTTAACTGCGATATTCTCGAGAATATCGCAGTTAAGCATTGTTTTG (R) AATTCAAAAACCAATGCTTAACTGCGATATTCTCGAGAATATCGCAGTTAAGCATTG
mDnmt2#2	(D) CCGGCCGCTAATTCTAGATAGTAATCTCGAGATTACTATCTAGAATTAGCGGTTTTG (R) AATTCAAAAACCGCTAATTCTAGATAGTAATCTCGAGATTACTATCTAGAATTAGCGG
mDnmt2#3	(D) CCGGCTTATAGAGTTCAGCTAAATTCTCGAGAATTTAGCTGAACCTCTATAAGTTTTG (R) AATTCAAAAACCTTATAGAGTTCAGCTAAATTCTCGAGAATTTAGCTGAACCTCTATAAG
hDNMT2#1	(D) CCGGCCAAAGTCATTGCTGCGATATCTCGAGATATCGCAGCAATGACTTTGGTTTTG (R) AATTCAAAAACCAAGTCATTGCTGCGATATCTCGAGATATCGCAGCAATGACTTTGG
hDNMT2#2	(D) CCGGAGCGTTTATAACACCATTATTCTCGAGAATAATGGTGTATAAACGCTTTTTG (R) AATTCAAAAACGCTTTTATAACACCATTATTCTCGAGAATAATGGTGTATAAACGCT
Luciferasa	(D) CCGGCCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGGGCGACTTAACCTTAGGTTTTG (R) AATTCAAAAACCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGGGCGACTTAACCTTAGG

Tabla suplementaria 6. Cebadores para la amplificación de regiones codificantes.

Clonación en vectores PyggyBac		
Gen	Secuencia directa 5'→3'	Secuencia inversa 5'→3'
mDnmt2	GCTAGCGAACCTCTGCGTGCTCTGGA	TTAATTAATCAAGATAGAATCTTGAATTACTAT
hDNMT2	GCTAGCGAGCCCTGCGGGTGC	TTAATTAATTATTCATATAAGATTTTGATTAGTT
Clonación en vectores FUW-Tet-On		
Gen	Secuencia directa 5'→3'	Secuencia inversa 5'→3'

<i>mDnmt2</i>	GAATTCATGGACTACAAAGAC	GAATTCTCAAGATAGAATCTTGG
<i>hDNMT2</i>	GAATTCATGGACTACAAAGACC	GAATTCATTATCATATAAGATTTTGATTAGTTT

Genotipado de clones *Dnmt2* KO

Oligonucleótido 1	Oligonucleótido 2	Oligonucleótido 3 (sgRNA)
GAGACGCAGAGAACCGGTAG	TAGTTTCATTGGCTGCACA	GAGCTGTACAGTGGCATTGG

Genotipado de clones *Dnmt2* KI

Secuencia directa 5'→3'	Secuencia inversa 5'→3'
CTGTCAAGAATGCCGCTAA	CCTGCTGAATTACCCAGGAA

Tabla suplementaria 7. Cebadores para la generación del donador pFETCH.

Región	Secuencia directa 5'→3'	Secuencia inversa 5'→3'
Región 1	TCCCCGACCTGCAGCCAGCTTCAATCCTCAAATGTATGAACCA	CCGGAACCTCCTCCGCTCCAGATAGAATCTTGGAACTACTATC
Región 2	AGTTCTTCTGATTGGAACATCTTCAGCTAAATCTTTGAATAGCA	TGGAGAGGACTTTCCAAGTACTGGGGCCATAACCTGT

Tabla suplementaria 8. Cebadores para el análisis mediante RT-qPCR.

Gen	Secuencia directa 5'→3'	Secuencia inversa 5'→3'
<i>Cdh1</i>	AACAACTGCATGAAGGCGGGAATC	CCTGTGCAGCTGGCTCAAATCAAA
<i>Cdh2</i>	GGCTGAAAATAGACCCCGTG	GCTGTGATGTTAATTGAGTTGGG
<i>Cdx2</i>	CAAGGACGTGAGCATGTATCC	GTAACCACCGTAGTCCGGGTGA
<i>Cldn3</i>	CCGGTCAGATGCAGTGCAAAATG	TGAGCAGAGCCGCAACAGGAAA
<i>Cxcl12</i>	TTTCAGATGCTTGACGTTGG	GCGCTCTGCATCAGTGAC
<i>Dnmt2</i> (m)	CTCTGCGTGTCTGGAGCT	TGCTTGATACCTTCATTAGCAACAGT
<i>DNMT2</i> (h)	CGGGTGCTGGAGCTATACAG	CGACAGTGTGACATCAATGGC
<i>Eomes</i>	CCAAAACACGGATATCACCC	GACCTCCAGGACAAATCTGA
<i>Epcam</i>	GCTGGCAACAAGTTGCTCTCTGAA	CGTTGCACTGCTTGGCTTTGAAGA
<i>Fgf5</i>	CTCAGGGGATTGTAGGAATACGAGGA	GGATCGCGGACGCATAGGTATTATAG
<i>Fgf8</i>	AATCGGACTCTGCTTCCAA	CAGGTCTGGCCAAACAAG
<i>Gata4</i>	CCCTACCCAGCCTACATGG	ACATATCGAGATTGGGGTGTCT
<i>Gata6</i>	TTGCTCCGTTAACAGCAGTG	GTGGTCTGCTGTGTAGAAGGA
<i>Hexa</i>	TGGCCCCAGTACATCCAAAC	GGTTACGGTAGCGTCGAAAGG
<i>Hoxd13</i>	GAAATCATCTTTCCAGGAGATG	CGCCGCTTGCTCTGTAAATG
<i>Hoxd3</i>	AAGCAGAAGAAGCAGTGTGC	TAGCGGTTGAAGTGGAATCC
<i>Hoxd4</i>	TGCACGTGAATTCGGTGAAC	GTGAGCGATTTCATCCGACG
<i>Hprt</i>	TCAGTCAACGGGGACATAAA	GGGGCTGTACTGCTTAACCAAG
<i>Map1b</i>	TAACCGAAGACACCTGAGG	TTGAGTCTCTGTTCCAAAGTTG
<i>Mdm2</i> (m)	GGATCTTGACGATGGCGTAAAG	AGGCTGTAAATCTCCGAGTCC
<i>MDM2</i> (h)	CCCCTTCCATCACATTGCAA	GCCCTCTTCAGCTTGTGTG
<i>Mixl1</i>	ATCCGCCCGGACCCTCCAAA	TCGGTCTGGAACACACCTGGA
<i>Nanog</i>	TTGCTCTTTCTGTGGAAAGG	CCAGGAAGACCCACACTCAT
<i>Nestin</i>	GAGTCTCAGCTCCAGCAG	AGATCGCTCAGATCTGGAA
<i>Ocln</i>	CCTCCAATGGCAAAGTGAATGGCA	TGTTTCATAGTGGTCAGGGTCCGT
<i>Oct4^{end}</i>	TCTTCCACCAGGCCCGGCTC	TGCGGGCGGACATGGGGAGATCC
<i>Pax3</i>	GCAGCGCAGGAGCAGAACCA	GCACTCGGGCTCGGTAAGC
<i>Slug</i>	CACATTCGAACCCACACATTGCCT	TGTGCCCTCAGGTTTGATCTGTCT
<i>Smyd1</i>	TCAGTGACCAGAGAGGGCTAC	AGCTCAATCTTGCCATTGTTGAA
<i>Sox17</i>	CGCACGGAATTCGAACAGTA	GTCAAATGTCGGGGTAGTTG
<i>T</i>	GCTTCAAGGAGCTAACTAACGAG	CCAGCAAGAAAGAGTACATGGC
<i>Trp53</i>	CTCTCCCCGCAAAAGAAAAA	CGGAACATCTCGAAGCGTTTA
<i>Twist1</i>	GGACAAGCTGAGCAAGATTCA	CGGAGAAGGCGTAGCTGAG
<i>Twist2</i>	CGCTACAGCAAGAAATCGAGC	GCTGAGCTTGTGAGAGGGG
<i>U6</i>	CTCGCTTCGGCAGCACATATACT	ACGCTTCACGAATTTGCGTGTC
<i>Vim</i>	CGTCCACACGCACCTACAG	GGGGGATGAGGAATAGAGGCT
<i>Wnt7b</i>	TTTGGGCTCCTCTACGTGAAG	CCCCGATCACAATGATGGCA
<i>Zeb1</i>	TGCTCACCTGCCGTATTGTGATA	AGTGCACTTGAACCTGCGGTTTCC

Tabla suplementaria 9. Cebadores para la síntesis de ADNc específico

Gen	Secuencia directa 5'→3'
<i>tRNA^{Asp}</i>	CTCAACTGGTGTCTGGAGTCGGCAATTCAGTTGAGTGGCTCCCCG
<i>BS-tRNA^{Asp}</i>	CTCAACTGGTGTCTGGAGTCGGCAATTCAGTTGAGTGGCTCCCCA

Tabla suplementaria 10. Cebadores para el análisis por PCR

Gen	Secuencia directa 5'→3'	Secuencia inversa 5'→3'
<i>Dnmt2</i>	GAACCTCTGCGTGTCTGGA	TCAAGATAGAATCTTGGAATTACTAT
<i>tRNA^{Asp}</i>	TAGTATAGTGGTGAGTATC	CGGCAATTCAGTTGAG
<i>BS-tRNA^{Asp}</i>	TAGTATAGTGGTGAGTATT	CGGCAATTCAGTTGAG

Tabla suplementaria 11. Cebadores para secuenciación por Sanger

Gen / Vector	Secuencia directa 5'→3'	Secuencia inversa 5'→3'
<i>Dnmt2</i>	TGCTTGCCCTTTTAGGATTG	CACCAAATTAGGAAAGGCACA
<i>BS-tRNA^{Asp}</i>	ACAGAGGGAAACCCCATCTT	n/a
pJET1.2	CGACTCACTATAGGAGAGCGGC	AAGAACATCGATTTCATGGCAG
SpCas9-2A-Puro V2.0	GAGGGCTATTTCATGAT	n/a
pLKO.1	CAAGGCTGTAGAGAGATAATTGGA	n/a
pFETCH Reg Homóloga 1	ACGCCTGTGAAACCGTACTA	CAGCAGGCTGAAGTTAGTAGC
pFETCH Reg Homóloga 2	GGCCGCTTTTCTGATTTCAT	GACGGCCAGTGAATTGGAG

9.3 Anticuerpos

Tabla suplementaria 12. Anticuerpos.

Anticuerpo	Aplicación / Concentración	Referencia
α-ACTB monoclonal de ratón	WB: 1/2.000	Santa Cruz Biotechnology, sc-47778
α-βTUBULINA monoclonal de ratón	WB: 1/800	Santa Cruz Biotechnology, sc-55529
α-CD49f monoclonal de ratón	FACS: 1/1.000	Invitrogen, 25-0495-82
α-CDH1 monoclonal de ratón	WB: 1/2.000	BD Biosciences, 610182
α-CDH2 monoclonal de ratón	WB: 1/2.000	Santa Cruz Biotechnology, sc-393933
α-DNMT2 monoclonal de ratón	WB: 1/1.000	Abcam, ab220175
α-FLAG-M2 monoclonal de ratón	WB: 1/1.000 IP: 2.5 µg/ mg proteína	Sigma-Aldrich, F1804
α-GAPDH policlonal de conejo	WB: 1/10.000	ProteinTech, 10494-1-AP
α-NANOG policlonal de conejo	WB: 1/1.000	Bethyl Laboratories, A300-397A-2
α-OCT4 monoclonal de ratón	WB: 1/2.000 IF: 1/1.000	Santa Cruz Biotechnology, sc-5279
α-P53 monoclonal de ratón	WB: 1/1.000	Santa Cruz Biotechnology, sc-126
α-SOX2 monoclonal de ratón	IF: 1/1.000	Santa Cruz Biotechnology, sc-365823
α-CONEJO IgG (H+L) HRP policlonal de asno	WB: 1/5.000	Invitrogen, A16023
α-RATÓN IgG (H+L) HRP policlonal de asno	WB: 1/5.000	Invitrogen, A16011
α-RATÓN TrueBlot Ultra Ig HRP monoclonal de rata	WB: 1/5.000	Rockland, 18-8817-31
α-RATÓN IgG (H+L) Alexa Fluor 488 policlonal de asno	IF: 1/500	Jackson Immuno Research, 711-545-150
α-RATÓN IgG (H+L) Rhodamine Red-X policlonal de asno	IF: 1/500	Jackson Immuno Research, 715-295-151

9.4 Conjuntos de datos externos

Tabla suplementaria 13. Conjuntos de datos externos

Descripción	Número acceso	Referencia
RNA-seq de MEFs, ESCs y iPSCs	GSE53212	(Carter et al., 2014)
RNA-seq durante SCR	GSE76233	(Cieply et al., 2016)
RNA BS-seq de ESCs	GSE83432	(Amort et al., 2017)
BS-RNA-seq de ESCs <i>Dnmt2</i> WT y KO	GSE81825	(Legrand et al., 2017)

9.5 Ontologías génicas y vías de señalización KEGG

Tabla suplementaria 14. Ontologías génicas.

Procesos biológicos	
Descripción	Código
Desarrollo embrionario uterino	GO:0001701
Desarrollo epitelial	GO:0060429
Diferenciación celular	GO:0030154
Morfogénesis celular en diferenciación	GO:0000904
Procesamiento del ARNm	GO:0006397
Proceso metabólico	GO:0008152
Regulación del metabolismo del ARN	GO:0051252
Traducción	GO:0006412
Funciones moleculares	
Descripción	Código
Unión a cromatina	GO:0003682
Unión del ARN	GO:0003723
Unión de proteínas	GO:0005515
Unión de proteínas quinasa	GO:0019901
Componentes celulares	
Descripción	Código
Citoplasma	GO:0005737
Cromatina	GO:0000785
Núcleo	GO:0005634
Nucleoplasma	GO:0005654

Tabla suplementaria 15. Vías de señalización KEGG.

Descripción	Código
Vía de señalización de P53	mmu04115
Vías en cáncer	mmu05200
Vías de señalización que regulan pluripotencia	mmu04550

10. LISTA DE ABREVIATURAS

5-AZA: 5-azacitidina.

ADN: ácido desoxirribo-nucleico.

ADNc: ADN complementario.

ANG: angiogenina.

AP: fosfatasa alcalina. Del inglés, *alkaline phosphatase*.

ARN: ácido ribonucleico.

ARNe: ARN derivado de *enhancers*.

ARNm: ARN mensajero.

ARNnc: ARN no codificante.

ARNlnc: ARN largo no codificante.

ARNr: ARN ribosómico.

ARNt: ARN transferente.

Asp: ácido aspártico.

BF: campo claro. Del inglés, *brightfield*.

BSA: albúmina de suero bovino. Del inglés *bovine serum albumin*.

RNA BS-seq: secuenciación masiva de ARN convertido con bisulfito. Del inglés, *RNA Bisulfite Sequencing*.

CDS: secuencia codificante. Del inglés, *coding sequence*.

CNV: variación del número de copias. Del inglés, *copy number variation*.

DAPI: 4',6-diamidino-2-fenilindol.

DAVID: base de datos para anotación, visualización y descubrimiento integrado. Del inglés, *database for annotation, visualization, and integrated discovery*.

DMEM: medio Eagle modificado de Dulbecco. Del inglés, *Dulbecco's modified medium*.

DMSO: dimetilsulfóxido.

DNMT(s): ADN metiltransferasa(s). Del inglés, *DNA methyltransferase (s)*.

DOX: doxiciclina.

DPBS: tampón fosfato salino de Dulbecco. Del inglés, *Dulbecco's phosphate buffered saline*.

DSB(s): rotura(s) de doble cadena. Del inglés, *double-strand breaks*,

EBs: cuerpos embrioides. Del inglés, *embryoid bodies*.

ELISA: ensayo de inmunoabsorción ligado a enzima. Del inglés, *enzyme-linked immunosorbent assay*

ER: receptor de estrógenos. Del inglés, *estrogen receptor*.

ESCs: células troncales embrionarias. Del inglés, *embryonic stem cells*.

EV: vector vacío. Del inglés, *empty vector*.

FACS: citometría de flujo activada por fluorescencia. Del inglés, *fluorescence-activated cell sorting*.

FBS: suero fetal bovino. Del inglés, *fetal bovine serum*.

FC: valor o tasa de cambio. Del inglés, *fold change*.

FDR: tasa de descubrimientos falsos. Del inglés, *false discovery rate*.

GFP: proteína fluorescente verde. Del inglés, *green fluorescent protein*.

Gly: glicina.

GO: ontología génica. Del inglés, *gene ontology*.

GSEA: análisis de enriquecimiento de conjuntos de genes. Del inglés *gene set enrichment analysis*.

GSK3β: glucógeno sintasa quinasa-3 beta.

hm⁵C: 5-hidroximetilcitosina (ARN).

HDFs: fibroblastos dérmicos humanos. Del inglés, *human dermal fibroblasts*.

HER2: receptor 2 del factor de crecimiento epidérmico humano. Del inglés, *human epidermal growth factor receptor 2*.

HR: recombinación homóloga. Del inglés, *homologous recombination*.

HZ: tasa o coeficiente de riesgo. Del inglés, *hazard ratio*.

ICM: masa celular interna. Del inglés, *inner cell mass*.

iHSCs: células troncales hematopoyéticas inducidas. Del inglés, *induced hematopoietic stem cell*.

IP: inmunoprecipitación.

iPSCs: células troncales pluripotentes inducidas. Del inglés, *induced pluripotent stem cells*.

KEGG: vías de señalización KEGG. Del inglés, *kyoto encyclopedia of genes and genomes*.

KI: introducción controlada o dirigida de manera específica de una secuencia de ADN particular en el genoma de un organismo. Del inglés *knockin*.

KO: nocaute, inactivación o supresión de un gen específico en el genoma de un organismo. Del inglés, *knockout*.

LB: medio Luria-Bertani.

LC-MS: cromatografía líquida-espectrometría de masas. Del inglés, *liquid chromatography-mass spectrometry*.

LIF: factor inhibidor de la leucemia. Del inglés, *leukemia inhibitory factor*.

m⁵C: modificación 5-metilcitosina en ARN. Del inglés, *5-methylcytosine*.

ME: mesendoderma.

MEFs: fibroblastos embrionarios murinos. Del inglés, *mouse embryonic fibroblasts*.

MEK: quinasa proteína-quinasa activada por mitógenos MEK. Del inglés, *mitogen-activated protein kinase kinase*.

MET: Transición mesenquimal a epitelial. Del inglés, *mesenchymal-to-epithelial transition*.

miRNA(s): micro ARN(s). Del inglés, *micro RNA(s)*.

MPO: ontología de fenotipos de mamíferos. Del inglés, *mammalian phenotype ontology*.

NE: neuroectoderma.

NES: valor de enriquecimiento normalizado. Del inglés, *normalized enrichment score*.

nLC-MS/MS: análisis de espectrometría de masas en tándem por cromatografía de nanolíquidos. Del inglés, *nano-liquid chromatography-mass spectrometry*.

NSUN: familia de proteínas NOL1/NOP2/sun.

OSKM: OCT4, SOX2, LKF4 y cMYC.

pb: pares de bases.

PB: PiggyBac

PC: componente principal. Del inglés, *principal component*.

PCA: análisis de componentes principales. Del inglés, *principal component analysis*.

PCR: reacción en cadena de la polimerasa. Del inglés, *polymerase chain reaction*.

PEI: polietilenimina.

pept.: péptidos.

PIC: cóctel inhibidor de proteasas. Del inglés, *protease inhibitor cocktail*.

PNK: polinucleótido quinasa.

PR: receptor de progesterona. Del inglés, *progesterone receptor*.

Pre-iPSCs: intermediarios avanzados de reprogramación.

PSM: coincidencia espectral de péptidos. Del inglés, *peptide-spectrum match*.

RT-qPCR: transcripción inversa seguida de reacción en cadena de la polimerasa cuantitativa. Del inglés, *reverse transcription-quantitative polymerase chain reaction*.

RA: ácido retinoico. Del inglés, *retinoic acid*.

RNA-seq: secuenciación masiva del ARN. Del inglés, *massive RNA sequencing*.

rpm: revoluciones por minuto.

RT: temperatura ambiente. Del inglés, *room temperature*.

SCR: reprogramación celular somática. Del inglés, *somatic cellular reprogramming*.

shRNA: ARN de horquilla corta. Del inglés, *short hairpin RNA*.

siRNA(s): ARN(s) pequeño(s) de interferencia. Del inglés, *small interfering RNA(s)*.

SL: medio de cultivo estándar pluripotente con presencia de suero y LIF.

TCGA: Atlas del Genoma del Cáncer. Del inglés, *The Cancer Genome Atlas*.

TE(s): elemento(s) transponible(s). Del inglés, *transposable element(s)*.

TET: enzimas metilcitosina dioxigenasas. Del inglés *ten-eleven translocation*.

TGF- β : factor de crecimiento transformante beta. Del inglés, *transforming growth factor β* .

tsRNAs: fragmentos derivados de ARNt. Del inglés, *transfer RNA-derived small RNAs*.

Val: valina.

WB: inmunotransferencia. Del inglés, *western blot*.

WH: cicatrización de herida. Del inglés, *wound healing*.

WT: salvaje. Del inglés, *wildtype*.

11. LISTA DE FIGURAS

Figura I1. Características principales de las ESCs.

Figura I2. Mecanismos reguladores del estado celular pluripotente.

Figura I3. Reprogramación celular somática mediada por OSKM.

Figura I4. Fases permisivas del proceso de SCR mediado por OSKM a partir de células somáticas de origen mesenquimal.

Figura I5. Las modificaciones epitranscriptómicas ocurren en diversas especies de ARN.

Figura I6. La modificación 5-metilcitosina se encuentra presente en moléculas de ADN y ARN.

Figura I7. La modificación m⁵C interviene en el metabolismo del ARN.

Figura I8. La deposición de m⁵C en ARN es catalizada por DNMT2 y la familia de proteínas NSUN.

Figura I9. La familia de proteínas DNMT.

Figura I10. DNMT2 cataliza la deposición de m⁵C en la citosina 38 de algunos ARNt.

Figura I11. Análisis de la divergencia y de la relación evolutiva de la proteína DNMT2.

Figura I12. La ausencia de m⁵C38 promueve la generación de tsRNAs.

Figura I13. La función de DNMT2 es crítica para la correcta síntesis de proteínas.

Figura I14. La actividad catalítica DNMT2 en ARN interviene en la herencia transgeneracional.

Figura R1. El silenciamiento de *Dnmt2* compromete la generación de iPSCs a partir de MEFs.

Figura R2. La falta de DNMT2 no afecta a la proliferación celular en MEFs.

Figura R3. La pérdida de función de DNMT2 impide la SCR a largo plazo.

Figura R4. La inhibición de *Dnmt2* reduce la activación de programas transcripcionales epiteliales.

Figura R5. DNMT2 interviene en la activación del programa transcripcional epitelial al inicio de la SCR.

Figura R6. El silenciamiento de *Dnmt2* afecta a la activación de *Cdh1* y a la represión de *Cdh2* en las primeras etapas de la SCR.

Figura R7. La pérdida de actividad catalítica de DNMT2 en ARN compromete la generación de iPSCs.

Figura R8. La ganancia de función de la actividad catalítica de DNMT2 favorece la generación de iPSCs.

Figura R9. La ausencia de DNMT2 no afecta a la reprogramación de intermediarios epiteliales a iPSCs.

Figura R10. La inhibición de DNMT2 dificulta la obtención de precursores hematopoyéticos a partir de MEFs.

Figura R11. La función de DNMT2 es necesaria para la generación eficiente de iHSCs humanas.

Figura R12. La sobreexpresión de *Dnmt2* facilita la obtención de precursores hematopoyéticos.

Figura R13. La ganancia de función de la actividad catalítica de DNMT2 favorece la generación de iHSCs humanas.

Figura R14. El silenciamiento de *Dnmt2* no altera la morfología de las ESCs.

Figura R15. La inhibición de *Dnmt2* no afecta la distribución de OCT4 y SOX2 en ESCs.

Figura R16. La pérdida de función de DNMT2 no altera la proliferación celular de las ESCs.

Figura R17. El silenciamiento de *Dnmt2* no afecta a la capacidad de autorrenovación de las ESCs.

Figura R18. Generación de una línea de ESCs KO para *Dnmt2*.

Figura R19. La ausencia de DNMT2 no compromete la capacidad de autorrenovación de las células pluripotentes.

Figura R20. La pérdida de DNMT2 no afecta a la proliferación de las ESCs.

Figura R21. Análisis de los transcriptomas de las ESCs WT y KO para *Dnmt2*.

Figura R22. La falta de DNMT2 altera el transcriptoma de las ESCs.

Figura R23. Efecto de la pérdida de DNMT2 en los niveles de la proteína P53 en ESCs.

Figura R24. La ausencia de DNMT2 afecta a la correcta activación de genes implicados en la diferenciación celular en ESCs.

Figura R25. Análisis del efecto de la ausencia de DNMT2 en los niveles de expresión de TEs en ESCs.

Figura R26. La presencia de DNMT2 es necesaria para la formación adecuada de EBs a partir de ESCs.

Figura R27. La ausencia de DNMT2 compromete la generación de gastruloides en cultivo.

Figura R28. La falta de DNMT2 dificulta la diferenciación hacia el linaje mesenquimal.

Figura R29. La ausencia de DNMT2 en ESCs no afecta a los niveles globales de m⁵C en ARN.

Figura R30. La falta de DNMT2 conduce a una pérdida específica de m⁵C38 en ARNt-Asp, ARNt-Gly y ARNt-Val.

Figura R31. Validación de la actividad catalítica de DNMT2 en la posición C38 del ARNt-Asp en ESCs.

Figura R32. Generación y validación de líneas de rescate de DNMT2 WT y mutantes catalíticos en ESCs KO para *Dnmt2*.

Figura R33. La pérdida de la actividad m⁵C-metiltransferasa de DNMT2 altera el perfil transcripcional de las ESCs.

Figura R34. DNMT2 interviene en la regulación de programas transcripcionales en ESCs mediante funciones dependientes e independientes de su actividad catalítica.

Figura R35. La actividad catalítica de DNMT2 es crítica en eventos de especificación temprana.

Figura R36. Identificación del interactoma de DNMT2 en ESCs.

Figura R37. Evaluación del impacto de la falta de DNMT2 en la expresión de genes que codifican sus proteínas interactoras.

Figura R38. Análisis de los procesos y fenotipos asociados a los interactores de DNMT2.

Figura R39. El gen *DNMT2* se encuentra mutado en una amplia variedad de tipos de cáncer.

Figura R40. Análisis de los genes con mutaciones conjuntas a *DNMT2* en pacientes con cáncer.

Figura R41. Asociación entre las mutaciones o la reducción de los niveles de *DNMT2* y el pronóstico desfavorable en pacientes con cáncer.

Figura R42. Evaluación de la expresión de *DNMT2* en diversos tipos de cáncer.

Figura R43. Asociación entre bajos niveles de *DNMT2* y una menor supervivencia en pacientes con cáncer de mama de los subtipos luminal A y luminal B.

Figura R44. La inhibición de *DNMT2* incrementa la capacidad de formación de colonias en la línea tumoral MCF-7.

Figura R45. El silenciamiento de *DNMT2* incrementa la capacidad de migración en la línea tumoral MCF-7.

Figura R46. La inhibición de *DNMT2* resulta en una disminución de los niveles de la proteína P53 en la línea tumoral MCF-7.

Figura D1. Participación de DNMT2 en la SCR mediante su actividad catalítica.

Figura D2. Potencial mecanismo según el que DNMT2 promueve la especificación del linaje mediante la metilación de ARNt.

Figura D3. Potencial mecanismo mediante el cual el silenciamiento de *DNMT2* conlleva un fenotipo más agresivo de la línea tumoral MCF-7 de cáncer de mama.

Tabla 1. DNMT2 y la familia de proteínas NSUN catalizan la deposición de m⁵C en diferentes sustratos de ARN.

Tabla 2. Fenotipos asociados a la pérdida de función de DNMT2.

12. BIBLIOGRAFÍA

- Adamczyk-Grochala, J., Bloniarz, D., Zielinska, K., Lewinska, A., y Wnuk, M. (2023). DNMT2/TRDMT1 gene knockout compromises doxorubicin-induced unfolded protein response and sensitizes cancer cells to ER stress-induced apoptosis. *Apoptosis* 28, 166-185.
- Adorno, M., Cordenonsi, M., Montagner, M., Dupont, S., Wong, C., Hann, B., Solari, A., Bobisse, S., Rondina, M.B., Guzzardo, V., *et al.* (2009). A Mutant-p53/Smad complex opposes p63 to empower TGFbeta-induced metastasis. *Cell* 137, 87-98.
- Aguilo, F., Li, S., Balasubramanian, N., Sancho, A., Benko, S., Zhang, F., Vashisht, A., Rengasamy, M., Andino, B., Chen, C.H., *et al.* (2016). Deposition of 5-Methylcytosine on Enhancer RNAs Enables the Coactivator Function of PGC-1alpha. *Cell Rep* 14, 479-492.
- Aguilo, F., Zhang, F., Sancho, A., Fidalgo, M., Di Cecilia, S., Vashisht, A., Lee, D.F., Chen, C.H., Rengasamy, M., Andino, B., *et al.* (2015). Coordination of m(6)A mRNA Methylation and Gene Transcription by ZFP217 Regulates Pluripotency and Reprogramming. *Cell Stem Cell* 17, 689-704.
- Amort, T., Rieder, D., Wille, A., Khokhlova-Cubberley, D., Riml, C., Trixl, L., Jia, X.Y., Micura, R., y Lusser, A. (2017). Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain. *Genome Biol* 18, 1.
- Apostolou, E., y Stadtfeld, M. (2018). Cellular trajectories and molecular mechanisms of iPSC reprogramming. *Curr Opin Genet Dev* 52, 77-85.
- Arguello, A.E., Li, A., Sun, X., Eggert, T.W., Mairhofer, E., y Kleiner, R.E. (2022). Reactivity-dependent profiling of RNA 5-methylcytidine dioxygenases. *Nat Commun* 13, 4176.
- Atlasi, Y., y Stunnenberg, H.G. (2017). The interplay of epigenetic marks during stem cell differentiation and development. *Nat Rev Genet* 18, 643-658.
- Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., y Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* 17, 126-140.
- Baillie-Johnson, P., van den Brink, S.C., Balayo, T., Turner, D.A., y Martinez Arias, A. (2015). Generation of Aggregates of Mouse Embryonic Stem Cells that Show Symmetry Breaking, Polarization and Emergent Collective Behaviour In Vitro. *J Vis Exp* 105, 53252.
- Balatti, V., Nigita, G., Veneziano, D., Drusco, A., Stein, G.S., Messier, T.L., Farina, N.H., Lian, J.B., Tomasello, L., Liu, C.G., *et al.* (2017). tsRNA signatures in cancer. *Proc Natl Acad Sci U S A* 114, 8071-8076.
- Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., *et al.* (2014). m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* 15, 707-719.
- Beccari, L., Moris, N., Girgin, M., Turner, D.A., Baillie-Johnson, P., Cossy, A.C., Lutolf, M.P., Duboule, D., y Arias, A.M. (2018). Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* 562, 272-276.
- Becker, M., Müller, S., Nellen, W., Jurkowski, T.P., Jeltsch, A., y Ehrenhofer-Murray, A.E. (2012). Pmt1, a Dnmt2 homolog in *Schizosaccharomyces pombe*, mediates tRNA methylation in response to nutrient signaling. *Nucleic Acids Res* 40, 11648-11658.

- Berglind, H., Pawitan, Y., Kato, S., Ishioka, C., y Soussi, T. (2008). Analysis of p53 mutation status in human cancer cell lines: a paradigm for cell line cross-contamination. *Cancer Biol Ther* 7, 699-708.
- Betlej, G., Lewinska, A., Adamczyk-Grochala, J., Bloniarz, D., Rzeszutek, I., y Wnuk, M. (2022). Deficiency of TRDMT1 impairs exogenous RNA-based response and promotes retrotransposon activity during long-term culture of osteosarcoma cells. *Toxicol In Vitro* 80, 105323.
- Biegging, K.T., Mello, S.S., y Attardi, L.D. (2014). Unravelling mechanisms of p53-mediated tumour suppression. *Nat Rev Cancer* 14, 359-370.
- Blanchard, J.W., Xie, J., El-Mecharrafie, N., Gross, S., Lee, S., Lerner, R.A., y Baldwin, K.K. (2017). Replacing reprogramming factors with antibodies selected from combinatorial antibody libraries. *Nat Biotechnol* 35, 960-968.
- Blanco, S., Bandiera, R., Popis, M., Hussain, S., Lombard, P., Aleksic, J., Sajini, A., Tanna, H., Cortés-Garrido, R., Gkatza, N., *et al.* (2016). Stem cell function and stress response are controlled by protein synthesis. *Nature* 534, 335-340.
- Blanco, S., Dietmann, S., Flores, J.V., Hussain, S., Kutter, C., Humphreys, P., Lukk, M., Lombard, P., Treps, L., Popis, M., *et al.* (2014). Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. *EMBO J* 33, 2020-2039.
- Blanco, S., Kurowski, A., Nichols, J., Watt, F.M., Benitah, S.A., y Frye, M. (2011). The RNA-methyltransferase Misu (NSun2) poises epidermal stem cells to differentiate. *PLoS Genet* 7, e1002403.
- Blaze, J., Navickas, A., Phillips, H.L., Heissel, S., Plaza-Jennings, A., Miglani, S., Asgharian, H., Foo, M., Katanski, C.D., Watkins, C.P., *et al.* (2021). Neuronal Nsun2 deficiency produces tRNA epitranscriptomic alterations and proteomic shifts impacting synaptic signaling and behavior. *Nat Commun* 12, 4913.
- Bloniarz, D., Adamczyk-Grochala, J., Lewinska, A., y Wnuk, M. (2021). The lack of functional DNMT2/TRDMT1 gene modulates cancer cell responses during drug-induced senescence. *Aging* 13, 15833-15874.
- Boccaletto, P., Machnicka, M.A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T.K., de Crecy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A., *et al.* (2018). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res* 46, D303-D307.
- Boccaletto, P., Stefaniak, F., Ray, A., Cappannini, A., Mukherjee, S., Purta, E., Kurkowska, M., Shirvanizadeh, N., Destefanis, E., Groza, P., *et al.* (2022). MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res* 50, D231-D235.
- Boo, S.H., y Kim, Y.K. (2020). The emerging role of RNA modifications in the regulation of mRNA stability. *Exp Mol Med* 52, 400-408.
- Bornelöv, S., Selmi, T., Flad, S., Dietmann, S., y Frye, M. (2019). Codon usage optimization in pluripotent embryonic stem cells. *Genome Biol* 20, 119.
- Boward, B., Wu, T., y Dalton, S. (2016). Concise Review: Control of Cell Fate Through Cell Cycle and Pluripotency Networks. *Stem Cells* 34, 1427-1436.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., *et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947-956.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., y Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* 2, 151-159.

Brumbaugh, J., Di Stefano, B., y Hochedlinger, K. (2019). Reprogramming: identifying the mechanisms that safeguard cell identity. *Development* 146, 182170.

Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., van Oudenaarden, A., y Jaenisch, R. (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 150, 1209-1222.

Buganim, Y., Faddah, D.A., y Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. *Nat Rev Genet* 14, 427-439.

Byrne, J.A., Pedersen, D.A., Clepper, L.L., Nelson, M., Sanger, W.G., Gokhale, S., Wolf, D.P., y Mitalipov, S.M. (2007). Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature* 450, 497-502.

Carter, A.C., Davis-Dusenbery, B.N., Koszka, K., Ichida, J.K., y Eggan, K. (2014). Nanog-independent reprogramming to iPSCs with canonical factors. *Stem Cell Reports* 2, 119-126.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., y Smith, A. (2003). Functional Expression Cloning of Nanog, a Pluripotency Sustaining Factor in Embryonic Stem Cells. *Cell* 113, 643-655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., y Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* 450, 1230-1234.

Che, Y.H., Lee, H., y Kim, Y.J. (2022). New insights into the epitranscriptomic control of pluripotent stem cell fate. *Exp Mol Med* 54, 1643-1651.

Chelmicki, T., Roger, E., Teissandier, A., Dura, M., Bonneville, L., Rucli, S., Dossin, F., Fouassier, C., Lameiras, S., y Bourc'his, D. (2021). m(6)A RNA methylation regulates the fate of endogenous retroviruses. *Nature* 591, 312-316.

Chen, C., Liu, W., Guo, J., Liu, Y., Liu, X., Liu, J., Dou, X., Le, R., Huang, Y., Li, C., *et al.* (2021a). Nuclear m(6)A reader YTHDC1 regulates the scaffold function of LINE1 RNA in mouse ESCs and early embryos. *Protein & cell* 12, 455-474.

Chen, H., Yang, H., Zhu, X., Yadav, T., Ouyang, J., Truesdell, S.S., Tan, J., Wang, Y., Duan, M., Wei, L., *et al.* (2020). m(5)C modification of mRNA serves a DNA damage code to promote homologous recombination. *Nat Commun* 11, 2834.

Chen, L., Ding, J., Wang, B., Chen, X., Ying, X., Yu, Z., y Dong, P. (2021b). RNA methyltransferase NSUN2 promotes hypopharyngeal squamous cell carcinoma proliferation and migration by enhancing TEAD1 expression in an m(5)C-dependent manner. *Exp cell res* 404, 112664.

- Chen, Q., Yan, M., Cao, Z., Li, X., Zhang, Y., Shi, J., Feng, G.H., Peng, H., Zhang, X., Zhang, Y., *et al.* (2016). Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* 351, 397-400.
- Chen, T., Hao, Y.J., Zhang, Y., Li, M.M., Wang, M., Han, W., Wu, Y., Lv, Y., Hao, J., Wang, L., *et al.* (2015). m(6)A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. *Cell Stem Cell* 16, 289-301.
- Chen, X., Li, A., Sun, B.F., Yang, Y., Han, Y.N., Yuan, X., Chen, R.X., Wei, W.S., Liu, Y., Gao, C.C., *et al.* (2019). 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat Cell Biol* 21, 978-990.
- Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., *et al.* (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106-1117.
- Cheng, J.X., Chen, L., Li, Y., Cloe, A., Yue, M., Wei, J., Watanabe, K.A., Shammoo, J.M., Anastasi, J., Shen, Q.J., *et al.* (2018). RNA cytosine methylation and methyltransferases mediate chromatin organization and 5-azacytidine response and resistance in leukaemia. *Nat Commun* 9, 1163.
- Chew, J.L., Loh, Y.H., Zhang, W., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P., *et al.* (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* 25, 6031-6046.
- Chuong, E.B., Rumi, M.A., Soares, M.J., y Baker, J.C. (2013). Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat Genet* 45, 325-329.
- Cieply, B., Park, J.W., Nakauka-Ddamba, A., Bebee, T.W., Guo, Y., Shang, X., Lengner, C.J., Xing, Y., y Carstens, R.P. (2016). Multiphasic and Dynamic Changes in Alternative Splicing during Induction of Pluripotency Are Coordinated by Numerous RNA-Binding Proteins. *Cell Rep* 15, 247-255.
- Cohn y Volkin (1951). Nucleoside-5'-Phosphates from Ribonucleic Acid. *Nature* 167, 483-484.
- Concordet, J.P., y Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* 46, W242-W245.
- Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T.A., Shliaha, P.V., Fidalgo, M., Saunders, A., Lawrence, M., Dietmann, S., *et al.* (2013). NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* 495, 370-374.
- Covelo-Molares, H., Souto, Y., Fidalgo, M., y Guallar, D. (2022). A Simple, Rapid, and Cost-Effective Method for Loss-of-Function Assays in Pluripotent Cells. *Methods Mol Biol* 2520, 199-213.
- Dai, X., Cheng, H., Bai, Z., y Li, J. (2017). Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. *J Cancer* 8, 3131-3141.
- Daniel, M.G., Pereira, C.F., Bernitz, J.M., Lemischka, I.R., y Moore, K. (2016). Reprogramming Mouse Embryonic Fibroblasts with Transcription Factors to Induce a Hemogenic Program. *J Vis Exp* 118, 54372.
- David, L., y Polo, J.M. (2014). Phases of reprogramming. *Stem Cell Res* 12, 754-761.

De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., *et al.* (2015). Hallmarks of pluripotency. *Nature* 525, 469-478.

Delatte, B., Wang, F., Ngoc, L.V., Collignon, E., Bonvin, E., Deplus, R., Calonne, E., Hassabi, B., Putmans, P., Awe, S., *et al.* (2016). RNA biochemistry. Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science* 351, 282-285.

Delaunay, S., y Frye, M. (2019). RNA modifications regulating cell fate in cancer. *Nat Cell Biol* 21, 552-559.

Delaunay, S., Helm, M., y Frye, M. (2023). RNA modifications in physiology and disease: towards clinical applications. *Nat Rev Genet*. Publicado solo en web.

Delaunay, S., Pascual, G., Feng, B., Klann, K., Behm, M., Hotz-Wagenblatt, A., Richter, K., Zaoui, K., Herpel, E., Munch, C., *et al.* (2022). Mitochondrial RNA modifications shape metabolic plasticity in metastasis. *Nature* 607, 593-603.

Deng, W., Jacobson, E.C., Collier, A.J., y Plath, K. (2021). The transcription factor code in iPSC reprogramming. *Curr Opin Genet Dev* 70, 89-96.

Deniz, O., Frost, J.M., y Branco, M.R. (2019). Regulation of transposable elements by DNA modifications. *Nat Rev Genet* 20, 417-431.

Dev, R.R., Ganji, R., Singh, S.P., Mahalingam, S., Banerjee, S., y Khosla, S. (2017). Cytosine methylation by DNMT2 facilitates stability and survival of HIV-1 RNA in the host cell during infection. *Biochem J* 474, 2009-2026.

Dhahbi, J.M., Spindler, S.R., Atamna, H., Boffelli, D., y Martin, D.I. (2014). Deep Sequencing of Serum Small RNAs Identifies Patterns of 5' tRNA Half and YRNA Fragment Expression Associated with Breast Cancer. *Biomark cancer* 6, 37-47.

Di Stefano, B., Sardina, J.L., van Oevelen, C., Collombet, S., Kallin, E.M., Vicent, G.P., Lu, J., Thieffry, D., Beato, M., y Graf, T. (2014). C/EBPalpha poises B cells for rapid reprogramming into induced pluripotent stem cells. *Nature* 506, 235-239.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., y Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* 29, 15-21.

Doetschman, T.C., Eistetter, H., Katz, M., Schmidt, W., y Kemler, R. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morphol* 87, 27-45.

Dong, A., Yoder, J.A., Zhang, X., Zhou, L., Bestor, T.H., y Cheng, X. (2001). Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. *Nucleic Acids Res* 29, 439-448.

Du, E., Li, J., Sheng, F., Li, S., Zhu, J., Xu, Y., y Zhang, Z. (2020). A pan-cancer analysis reveals genetic alterations, molecular mechanisms, and clinical relevance of m(5)C regulators. *Clin Transl Med* 10, e180.

Durdevic, Z., Hanna, K., Gold, B., Pollex, T., Cherry, S., Lyko, F., y Schaefer, M. (2013a). Efficient RNA virus control in *Drosophila* requires the RNA methyltransferase Dnmt2. *EMBO Rep* 14, 269-275.

- Durdevic, Z., Mobin, M.B., Hanna, K., Lyko, F., y Schaefer, M. (2013b). The RNA methyltransferase Dnmt2 is required for efficient Dicer-2-dependent siRNA pathway activity in *Drosophila*. *Cell Rep* 4, 931-937.
- Durinck, S., Spellman, P.T., Birney, E., y Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 4, 1184-1191.
- Elhardt, W., Shanmugam, R., Jurkowski, T.P., y Jeltsch, A. (2015). Somatic cancer mutations in the DNMT2 tRNA methyltransferase alter its catalytic properties. *Biochimie* 112, 66-72.
- Emara, M.M., Ivanov, P., Hickman, T., Dawra, N., Tisdale, S., Kedersha, N., Hu, G.F., y Anderson, P. (2010). Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *J Biol Chem* 285, 10959-10968.
- Evans, M., J., y Kaufman, M., H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154-156.
- Falcicchio, M., Ward, J.A., Macip, S., y Doveston, R.G. (2020). Regulation of p53 by the 14-3-3 protein interaction network: new opportunities for drug discovery in cancer. *Cell Death Discov* 6, 126-147.
- Falconi, M., Giangrossi, M., Zabaleta, M.E., Wang, J., Gambini, V., Tilio, M., Bencardino, D., Occhipinti, S., Belletti, B., Laudadio, E., *et al.* (2019). A novel 3'-tRNA(Glu)-derived fragment acts as a tumor suppressor in breast cancer by targeting nucleolin. *Faseb j* 33, 13228-13240.
- Fatima, N., Saif Ur Rahman, M., Qasim, M., Ali Ashfaq, U., Ahmed, U., y Shareef Masoud, M. (2023). Transcriptional Factors mediated Reprogramming to Pluripotency. *Curr Stem Cell Res Ther*. Publicado solo en web.
- Felder, M., Kapur, A., Gonzalez-Bosquet, J., Horibata, S., Heintz, J., Albrecht, R., Fass, L., Kaur, J., Hu, K., Shojaei, H., *et al.* (2014). MUC16 (CA125): tumor biomarker to cancer therapy, a work in progress. *Mol Cancer* 13, 129.
- Feschotte, C., y Pritham, E., J. (2007). DNA Transposons and the Evolution of Eukaryotic Genomes. *Annu Rev Genet* 41, 331-368.
- Fisher, O., Siman-Tov, R., y Ankri, S. (2006). Pleiotropic phenotype in *Entamoeba histolytica* overexpressing DNA methyltransferase (EhMeth). *Mol Biochem Parasitol* 147, 48-54.
- Fong, H., Hohenstein, K.A., y Donovan, P.J. (2008). Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. *Stem Cells* 26, 1931-1938.
- Forbes, S.A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S., *et al.* (2015). COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res* 43, D805-D811.
- Friedli, M., Turelli, P., Kapopoulou, A., Rauwel, B., Castro-Díaz, N., Rowe, H.M., Ecco, G., Unzu, C., Planet, E., Lombardo, A., *et al.* (2014). Loss of transcriptional control over endogenous retroelements during reprogramming to pluripotency. *Genome Res* 24, 1251-1259.
- Frye, M., Dragoni, I., Chin, S.F., Spiteri, I., Kurowski, A., Provenzano, E., Green, A., Ellis, I.O., Grimmer, D., Teschendorff, A., *et al.* (2010). Genomic gain of 5p15 leads to over-expression of Misu (NSUN2) in breast cancer. *Cancer Lett* 289, 71-80.

- Frye, M., Harada, B., Behm, M., y He, C. (2018). RNA modifications modulate gene expression during development. *Science* 361, 1346-1349.
- Fu, H., Zhang, W., Li, N., Yang, J., Ye, X., Tian, C., Lu, X., y Liu, L. (2021). Elevated retrotransposon activity and genomic instability in primed pluripotent stem cells. *Genome Biol* 22, 201.
- Fu, L., Guerrero, C.R., Zhong, N., Amato, N.J., Liu, Y., Liu, S., Cai, Q., Ji, D., Jin, S.G., Niedernhofer, L.J., *et al.* (2014). Tet-mediated formation of 5-hydroxymethylcytosine in RNA. *J Am Chem Soc* 136, 11582-11585.
- Fuentes-Iglesias, A., Ameneiro, C., Guallar, D., y Fidalgo, M. (2022). Detecting and Modulating ER Stress to Improve Generation of Induced Pluripotent Stem Cells. *Methods Mol Biol* 2454, 743-754.
- Gao, S., Tao, L., Hou, X., Xu, Z., Liu, W., Zhao, K., Guo, M., Wang, H., Cai, T., Tian, J., *et al.* (2016). Genome-wide gene expression analyses reveal unique cellular characteristics related to the amenability of HPC/HSCs into high-quality induced pluripotent stem cells. *Stem Cell Res Ther* 7, 40.
- Gao, Y., y Fang, J. (2021). RNA 5-methylcytosine modification and its emerging role as an epitranscriptomic mark. *RNA Biol* 18, 117-127.
- Garcia-Outeiral, V., de la Parte, C., Fidalgo, M., y Guallar, D. (2020). The Complexity of TET2 Functions in Pluripotency and Development. *Front Cell Dev Biol* 8, 1-9.
- Garcia-Vilchez, R., Sevilla, A., y Blanco, S. (2019). Post-transcriptional regulation by cytosine-5 methylation of RNA. *Biochim Biophys Acta Gene Regul Mech* 1862, 240-252.
- Genencher, B., Durdevic, Z., Hanna, K., Zinkl, D., Mobin, M.B., Senturk, N., Da Silva, B., Legrand, C., Carre, C., Lyko, F., *et al.* (2018). Mutations in Cytosine-5 tRNA Methyltransferases Impact Mobile Element Expression and Genome Stability at Specific DNA Repeats. *Cell Rep* 22, 1861-1874.
- Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., *et al.* (2015). Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science* 347, 1002-1006.
- Ghanbarian, H., Wagner, N., Polo, B., Baudouy, D., Kiani, J., Michiels, J.F., Cuzin, F., Rassoulzadegan, M., y Wagner, K.D. (2016). Dnmt2/Trdm1 as Mediator of RNA Polymerase II Transcriptional Activity in Cardiac Growth. *PLoS One* 11, e0156953.
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, y Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat methods* 6, 343-345.
- Gingold, H., Tehler, D., Christoffersen, N.R., Nielsen, M.M., Asmar, F., Kooistra, S.M., Christophersen, N.S., Christensen, L.L., Borre, M., Sørensen, K.D., *et al.* (2014). A dual program for translation regulation in cellular proliferation and differentiation. *Cell* 158, 1281-1292.
- Girardini, J.E., Napoli, M., Piazza, S., Rustighi, A., Marotta, C., Radaelli, E., Capaci, V., Jordan, L., Quinlan, P., Thompson, A., *et al.* (2011). A Pin1/mutant p53 axis promotes aggressiveness in breast cancer. *Cancer Cell* 20, 79-91.
- Golipour, A., David, L., Liu, Y., Jayakumaran, G., Hirsch, C.L., Trcka, D., y Wrana, J.L. (2012). A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell* 11, 769-782.

- Goll, M., Kirpekar, F., Maggert, K., Yoder, J., Hsieh, C.-L., Zhang, X., Golic, K., Jacobsen, S., y Bestor, T. (2006). Methylation of tRNA^{Asp} by the DNA Methyltransferase Homolog Dnmt2. *Science* 311, 395-398.
- Goodarzi, H., Liu, X., Nguyen, H.C., Zhang, S., Fish, L., y Tavazoie, S.F. (2015). Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. *Cell* 161, 790-802.
- Guallar, D., Bi, X., Pardavila, J.A., Huang, X., Saenz, C., Shi, X., Zhou, H., Faiola, F., Ding, J., Haruehanroengra, P., *et al.* (2018). RNA-dependent chromatin targeting of TET2 for endogenous retrovirus control in pluripotent stem cells. *Nat Genet* 50, 443-451.
- Guallar, D., Fuentes-Iglesias, A., Souto, Y., Ameneiro, C., Freire-Agulleiro, O., Pardavila, J.A., Escudero, A., Garcia-Outeiral, V., Moreira, T., Saenz, C., *et al.* (2020). ADAR1-Dependent RNA Editing Promotes MET and iPSC Reprogramming by Alleviating ER Stress. *Cell Stem Cell* 27, 300-314.
- Guo, G., Pinello, L., Han, X., Lai, S., Shen, L., Lin, T.W., Zou, K., Yuan, G.C., y Orkin, S.H. (2016). Serum-Based Culture Conditions Provoke Gene Expression Variability in Mouse Embryonic Stem Cells as Revealed by Single-Cell Analysis. *Cell Rep* 14, 956-965.
- Gyorffy, B. (2021). Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. *Comput Struct Biotechnol J* 19, 4101-4109.
- Haag, S., Sloan, K.E., Ranjan, N., Warda, A.S., Kretschmer, J., Blessing, C., Hubner, B., Seikowski, J., Dennerlein, S., Rehling, P., *et al.* (2016). NSUN3 and ABH1 modify the wobble position of mt-tRNA^{Met} to expand codon recognition in mitochondrial translation. *EMBO J* 35, 2104-2119.
- Hackett, J.A., Kobayashi, T., Dietmann, S., y Surani, M.A. (2017). Activation of Lineage Regulators and Transposable Elements across a Pluripotent Spectrum. *Stem cell reports* 8, 1645-1658.
- Hafner, A., Bulyk, M.L., Jambhekar, A., y Lahav, G. (2019). The multiple mechanisms that regulate p53 activity and cell fate. *Nat Rev Mol Cell Biol* 20, 199-210.
- Hansson, J., Rafiee, M.R., Reiland, S., Polo, J.M., Gehring, J., Okawa, S., Huber, W., Hochedlinger, K., y Krijgsvelde, J. (2012). Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. *Cell Rep* 2, 1579-1592.
- Harbeck, N., Penault-Llorca, F., Cortes, J., Gnant, M., Houssami, N., Poortmans, P., Ruddy, K., Tsang, J., y Cardoso, F. (2019). Breast cancer. *Nat Rev Dis Primers* 5, 66.
- Hassani, S.N., Moradi, S., Taleahmad, S., Braun, T., y Baharvand, H. (2019). Transition of inner cell mass to embryonic stem cells: mechanisms, facts, and hypotheses. *Cell Mol Life Sci* 76, 873-892.
- Haussecker, D., Huang, Y., Lau, A., Parameswaran, P., Fire, A.Z., y Kay, M.A. (2010). Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* 16, 673-695.
- Hayashi, K., de Sousa Lopes, S.M.C., Tang, F., Lao, K., y Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* 3, 391-401.
- He, C., Bozler, J., Janssen, K.A., Wilusz, J.E., Garcia, B.A., Schorn, A.J., y Bonasio, R. (2021a). TET2 chemically modifies tRNAs and regulates tRNA fragment levels. *Nat Struct Mol Biol* 28, 62-70.

- He, J., Babarinde, I.A., Sun, L., Xu, S., Chen, R., Shi, J., Wei, Y., Li, Y., Ma, G., Zhuang, Q., *et al.* (2021b). Identifying transposable element expression dynamics and heterogeneity during development at the single-cell level with a processing pipeline scTE. *Nat Commun* 12, 1456-1470.
- He, Y., Yu, X., Li, J., Zhang, Q., Zheng, Q., y Guo, W. (2020). Role of m(5)C-related regulatory genes in the diagnosis and prognosis of hepatocellular carcinoma. *Am j transl res* 12, 912-922.
- Heissenberger, C., Liendl, L., Nagelreiter, F., Gonskikh, Y., Yang, G., Stelzer, E.M., Krammer, T.L., Micutkova, L., Vogt, S., Kreil, D.P., *et al.* (2019). Loss of the ribosomal RNA methyltransferase NSUN5 impairs global protein synthesis and normal growth. *Nucleic Acids Res* 47, 11807-11825.
- Heissenberger, C., Rollins, J.A., Krammer, T.L., Nagelreiter, F., Stocker, I., Wacheul, L., Shpylovyi, A., Tav, K., Snow, S., Grillari, J., *et al.* (2020). The ribosomal RNA m(5)C methyltransferase NSUN-1 modulates healthspan and oogenesis in *Caenorhabditis elegans*. *Elife* 9, e56205.
- Hermann, A., Schmitt, S., y Jeltsch, A. (2003). The Human Dnmt2 Has Residual DNA-(Cytosine-C5) Methyltransferase Activity. *J Biol Chem* 278, 31717-31721.
- Hochedlinger, K., y Jaenisch, R. (2015). Induced Pluripotency and Epigenetic Reprogramming. *Cold Spring Harb Perspect Biol* 7, a019448.
- Hoffding, M.K., y Hyttel, P. (2015). Ultrastructural visualization of the Mesenchymal-to-Epithelial Transition during reprogramming of human fibroblasts to induced pluripotent stem cells. *Stem Cell Res* 14, 39-53.
- Honda, S., Loher, P., Shigematsu, M., Palazzo, J.P., Suzuki, R., Imoto, I., Rigoutsos, I., y Kirino, Y. (2015). Sex hormone-dependent tRNA halves enhance cell proliferation in breast and prostate cancers. *Proc Natl Acad Sci U S A* 112, E3816-3825.
- Huang da, W., Sherman, B.T., y Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4, 44-57.
- Huang, T., Chen, W., Liu, J., Gu, N., y Zhang, R. (2019). Genome-wide identification of mRNA 5-methylcytosine in mammals. *Nat Struct Mol Biology* 26, 380-388.
- Huang, W., Lan, M.-D., Qi, C.-B., Zheng, S.-J., Wei, S.-Z., Yuan, B.-F., y Feng, Y.-Q. (2016). Formation and determination of the oxidation products of 5-methylcytosine in RNA. *Chem Sci* 7, 5495-5502.
- Huber, S.M., van Delft, P., Mendil, L., Bachman, M., Smollett, K., Werner, F., Miska, E.A., y Balasubramanian, S. (2015). Formation and abundance of 5-hydroxymethylcytosine in RNA. *Chembiochem* 16, 752-755.
- Hussain, S., Aleksic, J., Blanco, S., Dietmann, S., y Frye, M. (2013a). Characterizing 5-methylcytosine in the mammalian epitranscriptome. *Genome Biol* 14, 215.
- Hussain, S., Tuorto, F., Menon, S., Blanco, S., Cox, C., Flores, J.V., Watt, S., Kudo, N.R., Lyko, F., y Frye, M. (2013b). The mouse cytosine-5 RNA methyltransferase NSun2 is a component of the chromatoid body and required for testis differentiation. *Mol Cell Biol* 33, 1561-1570.
- Ignatova, V.V., Kaiser, S., Ho, J.S.Y., Bing, X., Stolz, P., Tan, Y.X., Lee, C.L., Gay, F.P.H., Lastres, P.R., Gerlini, R., *et al.* (2020a). METTL6 is a tRNA m(3)C methyltransferase that regulates pluripotency and tumor cell growth. *Sci Adv* 6, eaaz4551.

Ignatova, V.V., Stolz, P., Kaiser, S., Gustafsson, T.H., Lastres, P.R., Sanz-Moreno, A., Cho, Y.L., Amarie, O.V., Aguilar-Pimentel, A., Klein-Rodewald, T., *et al.* (2020b). The rRNA m(6)A methyltransferase METTL5 is involved in pluripotency and developmental programs. *Genes Dev* 34, 715-729.

Ivanov, P., Emara, M.M., Villen, J., Gygi, S.P., y Anderson, P. (2011). Angiogenin-induced tRNA fragments inhibit translation initiation. *Mol Cell* 43, 613-623.

Ivanov, P., O'Day, E., Emara, M.M., Wagner, G., Lieberman, J., y Anderson, P. (2014). G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. *Proc Natl Acad Sci U S A* 111, 18201-18206.

Jeltsch, A., Ehrenhofer-Murray, A., Jurkowski, T.P., Lyko, F., Reuter, G., Ankri, S., Nellen, W., Schaefer, M., y Helm, M. (2017). Mechanism and biological role of Dnmt2 in Nucleic Acid Methylation. *RNA Biol* 14, 1108-1123.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J., A., y Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337, 816-821.

Jonkhout, N., Tran, J., Smith, M.A., Schonrock, N., Mattick, J.S., y Novoa, E.M. (2017). The RNA modification landscape in human disease. *RNA* 23, 1754-1769.

Jurkowski, T.P., y Jeltsch, A. (2011). On the evolutionary origin of eukaryotic DNA methyltransferases and Dnmt2. *PLoS One* 6, e28104.

Jurkowski, T.P., Meusburger, M., Phalke, S., Helm, M., Nellen, W., Reuter, G., y Jeltsch, A. (2008). Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism. *RNA* 14, 1663-1670.

Kadumuri, R.V., y Janga, S.C. (2018). Epitranscriptomic Code and Its Alterations in Human Disease. *Trends Mol Med* 24, 886-903.

Kaiser, S., Jurkowski, T.P., Kellner, S., Schneider, D., Jeltsch, A., y Helm, M. (2017). The RNA methyltransferase Dnmt2 methylates DNA in the structural context of a tRNA. *RNA Biol* 14, 1241-1251.

Katter, K., Geurts, A.M., Hoffmann, O., Mates, L., Landa, V., Hiripi, L., Moreno, C., Lazar, J., Bashir, S., Zidek, V., *et al.* (2013). Transposon-mediated transgenesis, transgenic rescue, and tissue-specific gene expression in rodents and rabbits. *FASEB J* 27, 930-941.

Kazazian, H., H. (2004). Mobile Elements: Drivers of Genome Evolution. *Science* 303, 1626-1632.

Keskin, H., Shen, Y., Huang, F., Patel, M., Yang, T., Ashley, K., Mazin, A.V., y Storici, F. (2014). Transcript-RNA-templated DNA recombination and repair. *Nature* 515, 436-439.

Khan, Muzammil A., Rafiq, Muhammad A., Noor, A., Hussain, S., Flores, Joana V., Rupp, V., Vincent, Akshita K., Malli, R., Ali, G., Khan, Falak S., *et al.* (2012). Mutation in NSUN2, which Encodes an RNA Methyltransferase, Causes Autosomal-Recessive Intellectual Disability. *Am J Hum Genet* 90, 856-863.

Khoddami, V., y Cairns, B.R. (2013). Identification of direct targets and modified bases of RNA cytosine methyltransferases. *Nat Biotechnol* 31, 458-464.

- Khodeer, S., Klungland, A., y Dahl, J.A. (2022). ALKBH5 regulates somatic cell reprogramming in a phase-specific manner. *J Cell Sci* 135, jcs259824.
- Kiani, J., Grandjean, V., Liebers, R., Tuorto, F., Ghanbarian, H., Lyko, F., Cuzin, F., y Rassoulzadegan, M. (2013). RNA-mediated epigenetic heredity requires the cytosine methyltransferase Dnmt2. *PLoS Genet* 9, e1003498.
- Kim, H.K., Fuchs, G., Wang, S., Wei, W., Zhang, Y., Park, H., Roy-Chaudhuri, B., Li, P., Xu, J., Chu, K., *et al.* (2017). A transfer-RNA-derived small RNA regulates ribosome biogenesis. *Nature* 552, 57-62.
- Kim, J., Chu, J., Shen, X., Wang, J., y Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132, 1049-1061.
- Kinoshita, M., y Smith, A. (2018). Pluripotency Deconstructed. *Development, growth & differentiation* 60, 44-52.
- Klawitter, S., Fuchs, N.V., Upton, K.R., Muñoz-Lopez, M., Shukla, R., Wang, J., Garcia-Cañadas, M., Lopez-Ruiz, C., Gerhardt, D.J., Sebe, A., *et al.* (2016). Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. *Nat Commun* 7, 10286.
- Krishna, S., Yim, D.G., Lakshmanan, V., Tirumalai, V., Koh, J.L., Park, J.E., Cheong, J.K., Low, J.L., Lim, M.J., Sze, S.K., *et al.* (2019). Dynamic expression of tRNA-derived small RNAs define cellular states. *EMBO Rep* 20, e47789.
- Kumar, P., Anaya, J., Mudunuri, S.B., y Dutta, A. (2014). Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biol* 12, 78.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., y Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular biology and evolution* 35, 1547-1549.
- Kumari, K., Groza, P., y Aguilo, F. (2021). Regulatory roles of RNA modifications in breast cancer. *NAR Cancer* 3, zcab036.
- Kuscu, C., Kumar, P., Kiran, M., Su, Z., Malik, A., y Dutta, A. (2018). tRNA fragments (tRFs) guide Ago to regulate gene expression post-transcriptionally in a Dicer-independent manner. *RNA* 24, 1093-1105.
- Lan, J., Rajan, N., Bizet, M., Penning, A., Singh, N.K., Guallar, D., Calonne, E., Li Greci, A., Bonvin, E., Deplus, R., *et al.* (2020). Functional role of Tet-mediated RNA hydroxymethylcytosine in mouse ES cells and during differentiation. *Nat Commun* 11, 4956.
- Legrand, C., Tuorto, F., Hartmann, M., Liebers, R., Jacob, D., Helm, M., y Lyko, F. (2017). Statistically robust methylation calling for whole-transcriptome bisulfite sequencing reveals distinct methylation patterns for mouse RNAs. *Genome Res* 27, 1589-1596.
- Lengner, C.J., Camargo, F.D., Hochedlinger, K., Welstead, G.G., Zaidi, S., Gokhale, S., Scholer, H.R., Tomilin, A., y Jaenisch, R. (2007). Oct4 expression is not required for mouse somatic stem cell self-renewal. *Cell Stem Cell* 1, 403-415.
- Letunic, I., y Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49, W293-W296.

Levine, A.J. (2020). p53: 800 million years of evolution and 40 years of discovery. *Nat Rev Cancer* 20, 471-480.

Lewinska, A., Adamczyk-Grochala, J., Kwasniewicz, E., Dereowska, A., Semik, E., Zabek, T., y Wnuk, M. (2018). Reduced levels of methyltransferase DNMT2 sensitize human fibroblasts to oxidative stress and DNA damage that is accompanied by changes in proliferation-related miRNA expression. *Redox Biol* 14, 20-34.

Lewinska, A., Adamczyk-Grochala, J., Kwasniewicz, E., y Wnuk, M. (2017). Downregulation of methyltransferase Dnmt2 results in condition-dependent telomere shortening and senescence or apoptosis in mouse fibroblasts. *J Cell Physiol* 232, 3714-3726.

Lewis, C.J., Pan, T., y Kalsotra, A. (2017). RNA modifications and structures cooperate to guide RNA-protein interactions. *Nat Rev Mol Cell Biol* 18, 202-210.

Li, H., Zhu, D., Wu, J., Ma, Y., Cai, C., Chen, Y., Qin, M., y Dai, H. (2021). New substrates and determinants for tRNA recognition of RNA methyltransferase DNMT2/TRDMT1. *RNA Biol* 18, 2531-2545.

Li, L., Li, M., Zheng, J., Li, Z., y Chen, J. (2023). Knocking down NSUN5 inhibits the development of clear cell renal cell carcinoma by inhibiting the p53 pathway. *Aging* 15, 4757-4773.

Li, M., y Belmonte, J.C. (2017). Ground rules of the pluripotency gene regulatory network. *Nat Rev Genet* 18, 180-191.

Li, M., y Izpisua Belmonte, J.C. (2018). Deconstructing the pluripotency gene regulatory network. *Nat Cell Biol* 20, 382-392.

Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., He, W., Chen, J., Li, F., Zhuang, Q., *et al.* (2010). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* 7, 51-63.

Li, S., y Mason, C.E. (2014). The pivotal regulatory landscape of RNA modifications. *Annu Rev Genomics Hum Genet* 15, 127-150.

Liao, H., Gaur, A., McConie, H., Shekar, A., Wang, K., Chang, J.T., Breton, G., y Denicourt, C. (2022). Human NOP2/NSUN1 regulates ribosome biogenesis through non-catalytic complex formation with box C/D snoRNPs. *Nucleic Acids Res* 50, 10695-10716.

Lim, J.P., Shyamasundar, S., Gunaratne, J., Scully, O.J., Matsumoto, K., y Bay, B.H. (2017). YBX1 gene silencing inhibits migratory and invasive potential via CORO1C in breast cancer in vitro. *BMC cancer* 17, 201.

Lin, M.J., Tang, L.Y., Reddy, M.N., y Shen, C.K. (2005). DNA methyltransferase gene dDnmt2 and longevity of *Drosophila*. *Journal of Biological Chemistry* 280, 861-864.

Lin, Y., Zhang, J., Li, Y., Guo, W., Chen, L., Chen, M., Chen, X., Zhang, W., Jin, X., Jiang, M., *et al.* (2022). CTPS1 promotes malignant progression of triple-negative breast cancer with transcriptional activation by YBX1. *J Transl Med* 20, 17.

Liu, J., Dou, X., Chen, C., Chen, C., Liu, C., Xu, M.M., Zhao, S., Shen, B., Gao, Y., Han, D., *et al.* (2020a). N (6)-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* 367, 580-586.

- Liu, J., Gao, M., He, J., Wu, K., Lin, S., Jin, L., Chen, Y., Liu, H., Shi, J., Wang, X., *et al.* (2021). The RNA m(6)A reader YTHDC1 silences retrotransposons and guards ES cell identity. *Nature* **591**, 322-326.
- Liu, J., Gao, M., Xu, S., Chen, Y., Wu, K., Liu, H., Wang, J., Yang, X., Wang, J., Liu, W., *et al.* (2020b). YTHDF2/3 Are Required for Somatic Reprogramming through Different RNA Deadenylation Pathways. *Cell Rep* **32**, 108120.
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., *et al.* (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol* **10**, 93-95.
- Liu, K., Cao, J., Shi, X., Wang, L., Y Zhao, T. (2020c). Cellular metabolism and homeostasis in pluripotency regulation. *Protein & cell* **11**, 630-640.
- Liu, T., Hu, X., Lin, C., Shi, X., He, Y., Zhang, J., y Cai, K. (2022). 5-methylcytosine RNA methylation regulators affect prognosis and tumor microenvironment in lung adenocarcinoma. *Ann Transl Med* **10**, 259.
- Liu, Y., y Santi, D.V. (2000). m5C RNA and m5C DNA methyl transferases use different cysteine residues as catalysts. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 8263-8265.
- Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., *et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* **38**, 431-440.
- Love, M.I., Huber, W., y Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550.
- Lu, W., y Kang, Y. (2019). Epithelial-Mesenchymal Plasticity in Cancer Progression and Metastasis. *Dev Cell* **49**, 361-374.
- Lucas, M.C., Pryszcz, L.P., Medina, R., Milenkovic, I., Camacho, N., Marchand, V., Motorin, Y., Ribas de Pouplana, L., y Novoa, E.M. (2023). Quantitative analysis of tRNA abundance and modifications by nanopore RNA sequencing. *Nat Biotechnol*. En línea antes de impresión.
- Lyko, F. (2018). The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat Rev Genet* **19**, 81-92.
- Lyu, H., Li, M., Jiang, Z., Liu, Z., y Wang, X. (2019). Correlate the TP53 Mutation and the HRAS Mutation with Immune Signatures in Head and Neck Squamous Cell Cancer. *Comput Struct Biotechnol J* **17**, 1020-1030.
- Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., y Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57-63.
- Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., *et al.* (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1**, 55-70.

Mao, J., Saiding, Q., Qian, S., Liu, Z., Zhao, B., Zhao, Q., Lu, B., Mao, X., Zhang, L., Zhang, Y., *et al.* (2022). Reprogramming stem cells in regenerative medicine. *1*, e20220005.

Marine, J.C., y Lozano, G. (2010). Mdm2-mediated ubiquitylation: p53 and beyond. *Cell Death Differ* *17*, 93-102.

Martin, G., R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* *78*, 7634-7638.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., *et al.* (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* *9*, 625-635.

Maute, R.L., Schneider, C., Sumazin, P., Holmes, A., Califano, A., Basso, K., y Dalla-Favera, R. (2013). tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc Natl Acad Sci U S A* *110*, 1404-1409.

Mei, L., Shen, C., Miao, R., Wang, J.Z., Cao, M.D., Zhang, Y.S., Shi, L.H., Zhao, G.H., Wang, M.H., Wu, L.S., *et al.* (2020). RNA methyltransferase NSUN2 promotes gastric cancer cell proliferation by repressing p57(Kip2) by an m(5)C-dependent manner. *Cell Death Dis* *11*, 270.

Metodiev, M.D., Spahr, H., Loguercio Polosa, P., Meharg, C., Becker, C., Altmueller, J., Habermann, B., Larsson, N.G., y Ruzzenente, B. (2014). NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly. *PLoS Genet* *10*, e1004110.

Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X., y Thomas, P.D. (2019). Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat Protoc* *14*, 703-721.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., y Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature* *454*, 49-55.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., y Yamanaka, S. (2003). The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells. *Cell* *113*, 631-642.

Morales-Hernández, A., González-Rico, F.J., Román, A.C., Rico-Leo, E., Alvarez-Barrientos, A., Sánchez, L., Macia, Á., Heras, S.R., García-Pérez, J.L., Merino, J.M., *et al.* (2016). Aluretrotransposons promote differentiation of human carcinoma cells through the aryl hydrocarbon receptor. *Nucleic Acids Res* *44*, 4665-4683.

Mori, S., Chang, J.T., Andrechek, E.R., Matsumura, N., Baba, T., Yao, G., Kim, J.W., Gatz, M., Murphy, S., y Nevins, J.R. (2009). Anchorage-independent cell growth signature identifies tumors with metastatic potential. *Oncogene* *28*, 2796-2805.

Motorin, Y., Lyko, F., y Helm, M. (2010). 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic Acids Res* *38*, 1415-1430.

Muller, P.A., y Vousden, K.H. (2014). Mutant p53 in cancer: new functions and therapeutic opportunities. *Cancer Cell* *25*, 304-317.

- Murakami, Y., Wei, F.Y., Kawamura, Y., Horiguchi, H., Kadomatsu, T., Miyata, K., Miura, K., Oike, Y., Ando, Y., Ueda, M., *et al.* (2023). NSUN3-mediated mitochondrial tRNA 5-formylcytidine modification is essential for embryonic development and respiratory complexes in mice. *Commun Biol* 6, 307.
- Nagy, Á., Munkácsy, G., y Györffy, B. (2021). Pancancer survival analysis of cancer hallmark genes. *Scientific reports* 11, 6047.
- Nakano, S., Suzuki, T., Kawarada, L., Iwata, H., Asano, K., y Suzuki, T. (2016). NSUN3 methylase initiates 5-formylcytidine biogenesis in human mitochondrial tRNA(Met). *Nat Chem Biol* 12, 546-551.
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., y Smith, A. (1998). Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. *Cell* 95, 379-391.
- Niwa, H., Miyazaki, J., y Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature* 24, 372-376.
- Nombela, P., Miguel-López, B., y Blanco, S. (2021). The role of m(6)A, m(5)C and Ψ RNA modifications in cancer: Novel therapeutic opportunities. *Mol Cancer* 20, 18.
- Ofusa, K., Chijimatsu, R., y Ishii, H. (2022). Detection techniques for epitranscriptomic marks. *Am J Physiol Cell Physiol* 322, C787-C793.
- Okano, M., Xie, S., y Li, E. (1998). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res* 26, 2536-2540.
- Okita, K., Ichisaka, T., y Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature* 448, 313-317.
- Orford, K.W., y Scadden, D.T. (2008). Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation. *Nat Rev Genet* 9, 115-128.
- Ortiz-Barahona, V., Soler, M., Davalos, V., Garcia-Prieto, C.A., Janin, M., Setien, F., Fernandez-Rebollo, I., Bech-Serra, J.J., De La Torre, C., Guil, S., *et al.* (2023). Epigenetic inactivation of the 5-methylcytosine RNA methyltransferase NSUN7 is associated with clinical outcome and therapeutic vulnerability in liver cancer. *Mol Cancer* 22, 83.
- Ou, M., Xu, X., Chen, Y., Li, L., Zhang, L., Liao, Y., Sun, W., Quach, C., Feng, J., y Tang, L. (2021). MDM2 induces EMT via the B-Raf signaling pathway through 14-3-3. *Oncol Rep* 46, 120.
- Parker, M.T., Knop, K., Sherwood, A.V., Schurch, N.J., Mackinnon, K., Gould, P.D., Hall, A.J., Barton, G.J., y Simpson, G.G. (2020). Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m(6)A modification. *Elife* 9, e49658.
- Pei, D., Shu, X., Gassama-Diagne, A., y Thiery, J.P. (2019). Mesenchymal-epithelial transition in development and reprogramming. *Nat Cell Biol* 21, 44-53.
- Pendse, S., Vaidya, A., y Kale, V. (2022). Clinical applications of pluripotent stem cells and their derivatives: current status and future perspectives. *Regenerative medicine* 17, 677-690.

- Pereira, C.F., Chang, B., Gomes, A., Bernitz, J., Papatsenko, D., Niu, X., Swiers, G., Azzoni, E., de Bruijn, M.F., Schaniel, C., *et al.* (2016). Hematopoietic Reprogramming In Vitro Informs In Vivo Identification of Hemogenic Precursors to Definitive Hematopoietic Stem Cells. *Dev Cell* 36, 525-539.
- Pereira, C.F., Chang, B., Qiu, J., Niu, X., Papatsenko, D., Hendry, C.E., Clark, N.R., Nomura-Kitabayashi, A., Kovacic, J.C., Ma'ayan, A., *et al.* (2013). Induction of a hemogenic program in mouse fibroblasts. *Cell Stem Cell* 13, 205-218.
- Perou, H., M., Sùrlie, T., Eisen, M., B., van de Rijn, M., Jeffrey, S., S., Rees, C., A., Pollack, J., R., Ross, D., T., Johnsen, H., Akslen, L., A., *et al.* (2000). Molecular portraits of human breast tumours. *Nature* 406, 747-752.
- Phalke, S., Nickel, O., Walluscheck, D., Hortig, F., Onorati, M.C., y Reuter, G. (2009). Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2. *Nat Genet* 41, 696-702.
- Plath, K., y Lowry, W.E. (2011). Progress in understanding reprogramming to the induced pluripotent state. *Nat Rev Genet* 12, 253-265.
- Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., *et al.* (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* 151, 1617-1632.
- Pontis, J., Pulver, C., Playfoot, C.J., Planet, E., Grun, D., Offner, S., Duc, J., Manfrin, A., Lutolf, M.P., y Trono, D. (2022). Primate-specific transposable elements shape transcriptional networks during human development. *Nat Commun* 13, 7178.
- Putri, G.H., Anders, S., Pyl, P.T., Pimanda, J.E., y Zanini, F. (2022). Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics* 38, 2943-2945.
- Radomska, H.S., Gonzalez, D.A., Okuno, Y., Iwasaki, H., Nagy, A., Akashi, K., Tenen, D.G., y Huettner, C.S. (2002). Transgenic targeting with regulatory elements of the human CD34 gene. *Blood* 100, 4410-4419.
- Rai, K., Chidester, S., Zavala, C.V., Manos, E.J., James, S.R., Karpf, A.R., Jones, D.A., y Cairns, B.R. (2007). Dnmt2 functions in the cytoplasm to promote liver, brain, and retina development in zebrafish. *Genes Dev* 21, 261-266.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., y Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8, 2281-2308.
- Robles, A.I., y Harris, C.C. (2010). Clinical outcomes and correlates of TP53 mutations and cancer. *Cold Spring Harb Perspect Biol* 2, a001016.
- Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H., y Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem* 280, 24731-24737.
- Rodriguez-Terrones, D., y Torres-Padilla, M.E. (2018). Nimble and Ready to Mingle: Transposon Outbursts of Early Development. *Trends Genet* 34, 806-820.
- Roundtree, I.A., Evans, M.E., Pan, T., y He, C. (2017). Dynamic RNA Modifications in Gene Expression Regulation. *Cell* 169, 1187-1200.

- Rowe, R.G., y Daley, G.Q. (2019). Induced pluripotent stem cells in disease modelling and drug discovery. *Nat Rev Genet* 20, 377-388.
- Ruiz, S., Panopoulos, A.D., Montserrat, N., Multon, M.C., Daury, A., Rocher, C., Spanakis, E., Batchelder, E.M., Orsini, C., Deleuze, J.F., *et al.* (2012). Generation of a drug-inducible reporter system to study cell reprogramming in human cells. *J Biol Chem* 287, 40767-40778.
- Sabapathy, K., y Lane, D.P. (2018). Therapeutic targeting of p53: all mutants are equal, but some mutants are more equal than others. *Nat Rev Clin Oncol* 15, 13-30.
- Sajini, A.A., Choudhury, N.R., Wagner, R.E., Bornelöv, S., Selmi, T., Spanos, C., Dietmann, S., Rappsilber, J., Michlewski, G., y Frye, M. (2019). Loss of 5-methylcytosine alters the biogenesis of vault-derived small RNAs to coordinate epidermal differentiation. *Nat Commun* 10, 2550.
- Sakashita, A., Kitano, T., Ishizu, H., Guo, Y., Masuda, H., Ariura, M., Murano, K., y Siomi, H. (2023). Transcription of MERVL retrotransposons is required for preimplantation embryo development. *Nat Genet* 55, 484-495.
- Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., y Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 7, 64-77.
- Sampath, P., Pritchard, D.K., Pabon, L., Reinecke, H., Schwartz, S.M., Morris, D.R., y Murry, C.E. (2008). A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* 2, 448-460.
- Savic, D., Partridge, E.C., Newberry, K.M., Smith, S.B., Meadows, S.K., Roberts, B.S., Mackiewicz, M., Mendenhall, E.M., y Myers, R.M. (2015). CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. *Genome Res* 25, 1581-1589.
- Schaefer, M., Hagemann, S., Hanna, K., y Lyko, F. (2009). Azacytidine inhibits RNA methylation at DNMT2 target sites in human cancer cell lines. *Cancer Res* 69, 8127-8132.
- Schaefer, M., Kapoor, U., y Jantsch, M.F. (2017). Understanding RNA modifications: the promises and technological bottlenecks of the 'epitranscriptome'. *Open Biol* 7, 170077.
- Schaefer, M., y Lyko, F. (2010). Lack of evidence for DNA methylation of Invader4 retroelements in *Drosophila* and implications for Dnmt2-mediated epigenetic regulation. *Nat Genet* 42, 920-921.
- Schaefer, M., Pollex, T., Hanna, K., Tuorto, F., Meusburger, M., Helm, M., y Lyko, F. (2010). RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev* 24, 1590-1595.
- Schaniel, C., y Moore, K.A. (2009). Genetic models to study quiescent stem cells and their niches. *Ann N Y Acad Sci* 1176, 26-35.
- Schimmel, P. (2018). The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. *Nat Rev Mol Cell Biol* 19, 45-58.
- Schorn, A.J., Gutbrod, M.J., LeBlanc, C., y Martienssen, R. (2017). LTR-Retrotransposon Control by tRNA-Derived Small RNAs. *Cell* 170, 61-71.

Selmi, T., Hussain, S., Dietmann, S., Heiss, M., Borland, K., Flad, S., Carter, J.M., Dennison, R., Huang, Y.L., Kellner, S., *et al.* (2021). Sequence- and structure-specific cytosine-5 mRNA methylation by NSUN6. *Nucleic Acids Res* 49, 1006-1022.

Shanmugam, R., Aklujkar, M., Schafer, M., Reinhardt, R., Nickel, O., Reuter, G., Lovley, D.R., Ehrenhofer-Murray, A., Nellen, W., Ankri, S., *et al.* (2014). The Dnmt2 RNA methyltransferase homolog of *Geobacter sulfurreducens* specifically methylates tRNA-Glu. *Nucleic Acids Res* 42, 6487-6496.

Shanmugam, R., Fierer, J., Kaiser, S., Helm, M., Jurkowski, T.P., y Jeltsch, A. (2015). Cytosine methylation of tRNA-Asp by DNMT2 has a role in translation of proteins containing poly-Asp sequences. *Cell Discov* 1, 15010.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., y Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498-2504.

Sharma, U., Conine, C.C., Shea, J.M., Boskovic, A., Derr, A.G., Bing, X.Y., Belleannee, C., Kucukural, A., Serra, R.W., Sun, F., *et al.* (2016). Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* 351, 391-396.

Sherman, B.T., Hao, M., Qiu, J., Jiao, X., Baseler, M.W., Lane, H.C., Imamichi, T., y Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* 50, W216-W221.

Shi, H., Chai, P., Jia, R., y Fan, X. (2020). Novel insight into the regulatory roles of diverse RNA modifications: Re-defining the bridge between transcription and translation. *Mol Cancer* 19, 78.

Shi, Y., Inoue, H., Wu, J.C., y Yamanaka, S. (2017). Induced pluripotent stem cell technology: a decade of progress. *Nature Rev Drug Discov* 16, 115-130.

Signer, R.A., Magee, J.A., Salic, A., y Morrison, S.J. (2014). Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature* 509, 49-54.

Silverio-Alves, R., Gomes, A.M., Kurochkin, I., Moore, K.A., y Pereira, C.F. (2019). Hemogenic Reprogramming of Human Fibroblasts by Enforced Expression of Transcription Factors. *J Vis Exp*, 154.

Smith, A., G., Heath, J., K., Donaldson, D., D., Wong, G., G., Moreau, J., Stahl, M., y Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* 336, 688-690.

Smith, Z.D., Nachman, I., Regev, A., y Meissner, A. (2010). Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat Biotechnol* 28, 521-526.

Smith, Z.D., Sindhu, C., y Meissner, A. (2016). Molecular features of cellular reprogramming and development. *Nat Rev Mol Cell Biol* 17, 139-154.

Sobala, A., y Hutvagner, G. (2013). Small RNAs derived from the 5' end of tRNA can inhibit protein translation in human cells. *RNA Biol* 10, 553-563.

Sommer, C.A., Sommer, A.G., Longmire, T.A., Christodoulou, C., Thomas, D.D., Gostissa, M., Alt, F.W., Murphy, G.J., Kotton, D.N., y Mostoslavsky, G. (2010). Excision of reprogramming transgenes

improves the differentiation potential of iPS cells generated with a single excisable vector. *Stem Cells* 28, 64-74.

Song, H., Zhang, J., Liu, B., Xu, J., Cai, B., Yang, H., Straube, J., Yu, X., y Ma, T. (2022). Biological roles of RNA m(5)C modification and its implications in Cancer immunotherapy. *Biomark Res* 10, 15.

Squires, J.E., Patel, H.R., Nousch, M., Sibbritt, T., Humphreys, D.T., Parker, B.J., Suter, C.M., y Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res* 40, 5023-5033.

Stadtfield, M., Maherali, N., Breault, D.T., y Hochedlinger, K. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2, 230-240.

Stefkova, K., Prochazkova, J., y Pachernik, J. (2015). Alkaline phosphatase in stem cells. *Stem Cells Int* 2015, 628368.

Su, D., Chan, C.T., Gu, C., Lim, K.S., Chionh, Y.H., McBee, M.E., Russell, B.S., Babu, I.R., Begley, T.J., y Dedon, P.C. (2014). Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. *Nat Protoc* 9, 828-841.

Sun, H.L., Zhu, A.C., Gao, Y., Terajima, H., Fei, Q., Liu, S., Zhang, L., Zhang, Z., Harada, B.T., He, Y.Y., *et al.* (2020a). Stabilization of ERK-Phosphorylated METTL3 by USP5 Increases m(6)A Methylation. *Mol Cell* 80, 633-647.

Sun, X., Huang, X., Lu, X., Wang, N., Wu, D., Yuan, M., Yin, H., y Wang, Q. (2021). The expression and clinical significance of the tRNA aspartic acid methyltransferase 1 protein in gastric cancer. *Int J Clin Oncol* 26, 2229-2236.

Sun, Z., Xue, S., Zhang, M., Xu, H., Hu, X., Chen, S., Liu, Y., Guo, M., y Cui, H. (2020b). Aberrant NSUN2-mediated m(5)C modification of H19 lncRNA is associated with poor differentiation of hepatocellular carcinoma. *Oncogene* 39, 6906-6919.

Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., *et al.* (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 49, D605-D612.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., y Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Current Biology* 11, 1553-1558.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., y Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861-872.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Sasaki, A., Yamamoto, M., Nakamura, M., Sutou, K., Osafune, K., y Yamanaka, S. (2014). Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. *Nat Commun* 5, 3678.

Takahashi, K., y Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.

Takahashi, K., y Yamanaka, S. (2015). A developmental framework for induced pluripotency. *Development* 142, 3274-3285.

- Tanabe, K., Nakamura, M., Narita, M., Takahashi, K., y Yamanaka, S. (2013). Maturation, not initiation, is the major roadblock during reprogramming toward pluripotency from human fibroblasts. *Proc Natl Acad Sci U S A* 110, 12172-12179.
- Tang, G., Cho, M., y Wang, X. (2022). OncoDB: an interactive online database for analysis of gene expression and viral infection in cancer. *Nucleic Acids Res* 50, D1334-D1339.
- Tang, Y., Xuan, Y., Qiao, G., Ou, Z., He, Z., Zhu, Q., Liao, M., y Yin, G. (2019). MDM2 promotes epithelial-mesenchymal transition through activation of Smad2/3 signaling pathway in lung adenocarcinoma. *OncoTargets Ther* 12, 2247-2258.
- Theunissen, T.W., van Oosten, A.L., Castelo-Branco, G., Hall, J., Smith, A., y Silva, J.C. (2011). Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. *Curr Biol* 21, 65-71.
- Thiagarajan, D., Dev, R.R., y Khosla, S. (2011). The DNA methyltransferase Dnmt2 participates in RNA processing during cellular stress. *Epigenetics* 6, 103-113.
- Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.P., y Mi, H. (2022). PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci* 31, 8-22.
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., y Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* 282, 1145-1147.
- Thomson, M., Liu, S.J., Zou, L.N., Smith, Z., Meissner, A., y Ramanathan, S. (2011). Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* 145, 875-889.
- Trixl, L., Amort, T., Wille, A., Zinni, M., Ebner, S., Hechenberger, C., Eichin, F., Gabriel, H., Schoberleitner, I., Huang, A., *et al.* (2018). RNA cytosine methyltransferase Nsun3 regulates embryonic stem cell differentiation by promoting mitochondrial activity. *Cell Mol Life Sci* 75, 1483-1497.
- Trixl, L., y Lusser, A. (2019). The dynamic RNA modification 5-methylcytosine and its emerging role as an epitranscriptomic mark. *WIREs RNA* 10, e1510.
- Trixl, L., Rieder, D., Amort, T., y Lusser, A. (2019). Bisulfite Sequencing of RNA for Transcriptome-Wide Detection of 5-Methylcytosine. *Methods Mol Biol* 1870, 1-21.
- Tuorto, F., Herbst, F., Alerasool, N., Bender, S., Popp, O., Federico, G., Reitter, S., Liebers, R., Stoecklin, G., Grone, H.J., *et al.* (2015). The tRNA methyltransferase Dnmt2 is required for accurate polypeptide synthesis during haematopoiesis. *EMBO J* 34, 2350-2362.
- Tuorto, F., Liebers, R., Musch, T., Schaefer, M., Hofmann, S., Kellner, S., Frye, M., Helm, M., Stoecklin, G., y Lyko, F. (2012). RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. *Nat Struct Mol Biol* 19, 900-905.
- Ungerleider, N.A., Rao, S.G., Shahbandi, A., Yee, D., Niu, T., Frey, W.D., y Jackson, J.G. (2018). Breast cancer survival predicted by TP53 mutation status differs markedly depending on treatment. *Breast Cancer Res* 20, 115.
- Urlinger, S., Baron, U., Thellmann, M., Hasan, M., Bujard, H., y Hillen, W. (2000). Exploring the sequence space for tetracyclinedependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci U S A* 97, 7963-7968.

Van Haute, L., Lee, S.Y., McCann, B.J., Powell, C.A., Bansal, D., Vasiliauskaitė, L., Garone, C., Shin, S., Kim, J.S., Frye, M., *et al.* (2019). NSUN2 introduces 5-methylcytosines in mammalian mitochondrial tRNAs. *Nucleic Acids Res* 47, 8720-8733.

Wakayama, T., Tabar, V., Rodriguez, I., Perry, A.C., Studer, L., y Mombaerts, P. (2001). Differentiation of embryonic stem cell lines generated from adult somatic cells by nuclear transfer. *Science* 292, 740-743.

Wang, H., Yang, Y., Liu, J., y Qian, L. (2021). Direct cell reprogramming: approaches, mechanisms and progress. *Nat Rev Mol Cell Biol* 22, 410-424.

Wang, J., Ma, G., Li, M., Han, X., Xu, J., Liang, M., Mao, X., Chen, X., Xia, T., Liu, X., *et al.* (2020a). Plasma tRNA Fragments Derived from 5' Ends as Novel Diagnostic Biomarkers for Early-Stage Breast Cancer. *Mol Ther Nucleic Acids* 21, 954-964.

Wang, W., Lin, C., Lu, D., Ning, Z., Cox, T., Melvin, D., Wang, X., Bradley, A., y Liu, P. (2008). Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc Natl Acad Sci U S A* 105, 9290-9295.

Wang, X.-H., Aliyari, R., Li, W.-X., Li, H.-W., Kim, K., Carthew, R., Atkinson, P., y Ding, S.-W. (2006). RNA Interference Directs Innate Immunity Against Viruses in Adult *Drosophila*. *Science* 312, 452-454.

Wang, X., y He, C. (2014). Dynamic RNA modifications in postranscriptional regulation. *Mol Cell* 56, 5-12.

Wang, Y., Li, D., Lu, J., Chen, L., Zhang, S., Qi, W., Li, W., y Xu, H. (2020b). Long noncoding RNA TTN-AS1 facilitates tumorigenesis and metastasis by maintaining TTN expression in skin cutaneous melanoma. *Cell Death Dis* 11, 664.

Wang, Y., Li, Y., Toth, J.I., Petroski, M.D., Zhang, Z., y Zhao, J.C. (2014). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat Cell Biol* 16, 191-198.

Wen, J., Lv, R., Ma, H., Shen, H., He, C., Wang, J., Jiao, F., Liu, H., Yang, P., Tan, L., *et al.* (2018). Zc3h13 Regulates Nuclear RNA m(6)A Methylation and Mouse Embryonic Stem Cell Self-Renewal. *Mol Cell* 69, 1028-1038.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., y Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448, 318-324.

Williams, R., L., Hiltont, D., J., Pease, S., Willson, T., A., Stewart, C., L., Gearing, D., P., Wagner, E., F., Metcalfe, D., Nicolat, N., A., y Gough, N., M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* 336, 684-687.

Wissing, S., Muñoz-Lopez, M., Macia, A., Yang, Z., Montano, M., Collins, W., Garcia-Perez, J.L., Moran, J.V., y Greene, W.C. (2012). Reprogramming somatic cells into iPS cells activates LINE-1 retroelement mobility. *Human Mol Genet* 21, 208-218.

Xi, C., Sun, J., Xu, X., Wu, Y., Kou, X., Zhao, Y., Shen, J., Dong, Y., Chen, K., Su, Z., *et al.* (2022). Methyl14-driven senescence-associated secretory phenotype facilitates somatic cell reprogramming. *Stem cell reports* 17, 1799-1809.

- Xing, M., Liu, Q., Mao, C., Zeng, H., Zhang, X., Zhao, S., Chen, L., Liu, M., Shen, B., Guo, X., *et al.* (2020). The 18S rRNA m(6) A methyltransferase METTL5 promotes mouse embryonic stem cell differentiation. *EMBO Rep* 21, e49863.
- Xue, S., Xu, H., Sun, Z., Shen, H., Chen, S., Ouyang, J., Zhou, Q., Hu, X., y Cui, H. (2019). Depletion of TRDMT1 affects 5-methylcytosine modification of mRNA and inhibits HEK293 cell proliferation and migration. *Biochem Biophys Res Commun* 520, 60-66.
- Yamanaka, S. (2020). Pluripotent Stem Cell-Based Cell Therapy-Promise and Challenges. *Cell Stem Cell* 27, 523-531.
- Yamanaka, S., y Blau, H.M. (2010). Nuclear reprogramming to a pluripotent state by three approaches. *Nature* 465, 704-712.
- Yang, H., Wang, Y., Xiang, Y., Yadav, T., Ouyang, J., Phoon, L., Zhu, X., Shi, Y., Zou, L., y Lan, L. (2022). FMRP promotes transcription-coupled homologous recombination via facilitating TET1-mediated m5C RNA modification demethylation. *Proc Natl Acad Sci U S A* 119, e2116251119.
- Yang, L., Perrera, V., Saplaoura, E., Apelt, F., Bahin, M., Kramdi, A., Olas, J., Mueller-Roeber, B., Sokolowska, E., Zhang, W., *et al.* (2019a). m(5)C Methylation Guides Systemic Transport of Messenger RNA over Graft Junctions in Plants. *Curr Biol* 29, 2465-2476.
- Yang, W.R., Ardeljan, D., Pacyna, C.N., Payer, L.M., y Burns, K.H. (2019b). SQuIRE reveals locus-specific regulation of interspersed repeat expression. *Nucleic Acids Res* 47, e27.
- Yang, X., Yang, Y., Sun, B.F., Chen, Y.S., Xu, J.W., Lai, W.Y., Li, A., Wang, X., Bhattarai, D.P., Xiao, W., *et al.* (2017). 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. *Cell Res* 27, 606-625.
- Yang, Y., Wang, L., Han, X., Yang, W.L., Zhang, M., Ma, H.L., Sun, B.F., Li, A., Xia, J., Chen, J., *et al.* (2019c). RNA 5-Methylcytosine Facilitates the Maternal-to-Zygotic Transition by Preventing Maternal mRNA Decay. *Mol Cell* 75, 1188-1202.
- Ying, Q.-L., Nichols, J., Evans, E.P., y Smith, A.G. (2002). Changing potency by spontaneous fusion. *Nature* 416, 545-548.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., *et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917-1920.
- Yu, T., Xie, Y., Tang, C., Wang, Y., Yuan, S., Zheng, H., y Yan, W. (2021). Dnmt2-null sperm block maternal transmission of a paramutant phenotypedagger. *Biol Reprod* 105, 603-612.
- Yu, W., Zhang, C., Wang, Y., Tian, X., Miao, Y., Meng, F., Ma, L., Zhang, X., y Xia, J. (2023). YAP 5-methylcytosine modification increases its mRNA stability and promotes the transcription of exosome secretion-related genes in lung adenocarcinoma. *Cancer Gene Ther* 30, 149-162.
- Zhang, H.-Y., Xiong, J., Qi, B.-L., Feng, Y.-Q., y Yuan, B.-F. (2016). The existence of 5-hydroxymethylcytosine and 5-formylcytosine in both DNA and RNA in mammals. *Chemical Communications* 52, 737-740.

Zhang, R., Gan, W., Zong, J., Hou, Y., Zhou, M., Yan, Z., Li, T., Lv, S., Zeng, Z., Wang, W., *et al.* (2023). Developing an m5C regulator-mediated RNA methylation modification signature to predict prognosis and immunotherapy efficacy in rectal cancer. *Front Immunol* 14, 1054700.

Zhang, Y., Zhang, X., Shi, J., Tuorto, F., Li, X., Liu, Y., Liebers, R., Zhang, L., Qu, Y., Qian, J., *et al.* (2018). Dnmt2 mediates intergenerational transmission of paternally acquired metabolic disorders through sperm small non-coding RNAs. *Nat Cell Biol* 20, 535-540.

Zhao, L., Qu, X., Wu, Z., Li, Y., Zhang, X., y Guo, W. (2020). TP53 somatic mutations are associated with poor survival in nonsmall cell lung cancer patients who undergo immunotherapy. *Aging* 12, 14556-14568.

Zhu, X., Wang, X., Yan, W., Yang, H., Xiang, Y., Lv, F., Shi, Y., Li, H.Y., y Lan, L. (2021). Ubiquitination-mediated degradation of TRDMT1 regulates homologous recombination and therapeutic response. *NAR Cancer* 3, zcab010.



Esta tesis doctoral investiga las funciones de DNMT2, tanto dependientes como independientes de su actividad catalítica en el ARN, durante la reprogramación celular, pluripotencia y cáncer. A través de diversos modelos y técnicas experimentales, se muestra su papel crucial en la adquisición y salida de la pluripotencia. Aunque prescindible para el mantenimiento del estado pluripotente, mutaciones y niveles reducidos de expresión de DNMT2 se asocian con un pronóstico desfavorable en distintos tipos de cáncer, aumentando la agresividad de líneas tumorales de cáncer de mama tras su silenciamiento. Este estudio revela un nuevo nivel de regulación controlado por las funciones de DNMT2, ofreciendo así nuevas perspectivas para comprender su participación más amplia en el desarrollo y en procesos patológicos.