



**UNIVERSIDAD
DE GRANADA**

Programa Oficial de Doctorado en Biomedicina

**Papel de TCERG1 en la integridad de los
cuerpos de Cajal y en la biogénesis de los
snRNPs**

Cristina Moreno Castro

IPBLN, CSIC

Granada 2020

Editor: Universidad de Granada. Tesis Doctorales
Autor: Cristina Moreno Castro
ISBN: 978-84-1117-826-6
URI: <https://hdl.handle.net/10481/81480>

La doctoranda CRISTINA MORENO CASTRO y el director CARLES M^a SUÑÉ
NEGRE de la tesis titulada **PAPEL DE TCERG1 EN LA INTEGRIDAD DE LOS
CUERPOS DE CAJAL Y EN LA BIOGÉNESIS DE LOS SNRNPS.**

Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por la doctoranda bajo la dirección del director de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 18 de Diciembre 2019

Director de la tesis

Doctoranda

Fdo: Carles M^a Suñé Negre

Fdo: Cristina Moreno Castro

Resumen

TCERG1 es una proteína implicada en la transcripción y el *splicing* de los mRNAs y en el acoplamiento funcional de ambos procesos. En esta tesis doctoral, hemos descubierto una nueva función para esta proteína: TCERG1 regula la homeostasis de los cuerpos de Cajal (CBs) y la biogénesis de las ribonucleoproteínas nucleares pequeñas (*small nuclear ribonucleoprotein*, snRNPs), partículas imprescindibles para el proceso de *splicing* del pre-mRNA.

Este estudio se inició con la identificación de la región de TCERG1 responsable de la unión con NOLC1, una fosfoproteína nucleolar componente de los CBs. Hemos identificado que los dominios FF4 y FF5 de TCERG1 son los responsables de la interacción entre TCERG1 y NOLC1 y que la inhibición de TCERG1 afecta a la localización de NOLC1 en los CBs nucleares pero no a su localización en el nucléolo. En relación a los CBs, mostramos que la inhibición de TCERG1 afecta la integridad de estos orgánulos ya que su proteína marcadora, la coilina, con la que interacciona TCERG1, se dispersa por el nucleoplasma en numerosos pequeños gránulos. Este efecto es específico de TCERG1, ya que los CBs pueden formarse de nuevo tras la reexpresión controlada de TCERG1. Además, la inhibición de TCERG1 no parece afectar de forma prominente a otros orgánulos nucleares como los *speckles*, *gems* o HLB. Paradójicamente, el efecto de

disgregación de los CBs puede ser observado también con la sobreexpresión saturante de TCERG1.

En esta tesis, hemos identificado también un papel importante de TCERG1 en la biogénesis de los snRNPs. Hemos observado que la inhibición de TCERG1 afecta la unión de las proteínas Sm a los RNAs pequeños nucleares ricos en uridinas (snRNAs), unión necesaria para formar el complejo snRNP maduro. La formación de este complejo tras la inhibición de TCERG1 se ve también disminuida al monitorizar la expresión de una proteína de fusión GFP-SmB. El papel de TCERG1 en la formación del complejo Sm-snRNP no parece implicar un efecto sobre la transcripción de los snRNAs dado que, en las condiciones experimentales usadas, los niveles de los snRNAs endógenos no están disminuidos. La interacción de TCERG1 con factores implicados en la biogénesis de los snRNPs, como CRM1, Int11 o SMN, refuerzan un posible papel de TCERG1 en este proceso.

Nuestros datos indican que TCERG1 juega un papel importante en la homeostasis de los CBs y en el ensamblaje de los complejos Sm-snRNAs durante la biogénesis de los snRNPs.

| | |
|---|----------|
| Introducción..... | 5 |
| 1. Arquitectura funcional del núcleo de la célula eucariota | 7 |
| 1.1. Formación de orgánulos nucleares | 9 |
| 1.2. Nucléolo..... | 10 |
| 1.2.1. NOLC1..... | 12 |
| 1.3. Cuerpos de Cajal (CBs, <i>Cajal Bodies</i>)..... | 13 |
| 1.3.1. Formación de los CBs | 17 |
| 1.4. <i>Gems</i> | 19 |
| 1.5. HLB (<i>histone locus body</i>) | 20 |
| 1.6. <i>Paraspeckles</i> | 21 |
| 1.7. PML-NB (<i>PML nuclear bodies</i>)..... | 22 |
| 1.8. <i>Speckles</i> nucleares..... | 23 |
| 1.8.1. Dominios de señalización hacia los <i>speckles</i> nucleares | 24 |
| 2. Biogénesis de los snRNPs | 26 |
| 2.1. Transcripción y procesamiento de los snRNAs | 27 |
| 2.2. Control de calidad y exporte de los snRNAs | 30 |
| 2.3. Ensamblaje citoplasmático de los snRNPs | 33 |
| 2.4. Importe nuclear de los snRNPs..... | 34 |
| 2.5. U6 snRNP | 35 |
| 2.6. Reciclaje de los snRNPs | 35 |
| 3. Proteínas con dominios WW y FF..... | 36 |
| 3.1. Características de los dominios WW y FF | 36 |
| 3.1.1. Dominios WW | 36 |
| 3.1.2. Dominios FF..... | 37 |
| 3.2. PRPF40A..... | 39 |
| 3.3. PRPF40B..... | 39 |

| | |
|---|-----------|
| 3.4. TCERG1, factor regulador de la elongación transcripcional (<i>Transcription Elongation Regulator 1</i>) | 40 |
| 3.4.1. Implicación de TCERG1 en el proceso de transcripción y <i>splicing</i> .. | 42 |
| 3.4.1. Localización de TCERG1 en los <i>speckles</i> nucleares..... | 43 |
| 3.4.2. Interacción de TCERG1 con NOLC1 | 44 |
| Objetivos | 47 |
| Material y métodos..... | 51 |
| 1. Plásmidos..... | 53 |
| 2. Cultivos celulares y transfecciones | 53 |
| 2.1. Cultivos celulares | 53 |
| 2.2. Transfecciones..... | 55 |
| 2.3. Generación de líneas establemente silenciadas mediante CRISPR/Cas9 | 56 |
| 3. Procedimientos de biología celular | 57 |
| 3.1. Inmunofluorescencia y microscopía confocal | 57 |
| 3.2. Sincronización del ciclo celular | 59 |
| 4. Manipulación y análisis de proteínas | 60 |
| 4.1. Preparación de extractos de proteínas | 61 |
| 4.1. Análisis por <i>western blot</i> | 61 |
| 4.2. Precipitación mediada por resina acoplada a GST (Ensayos de <i>pull down</i>) | 62 |
| 4.3. Tinción de plata | 63 |
| 4.4. Ensayos de inmunoprecipitación..... | 63 |
| 4.4.1. Inmunoprecipitación de proteínas endógenas | 63 |
| 4.4.2. Inmunoprecipitación de proteínas expresadas transitoriamente con el epítipo T7..... | 64 |
| 4.4.3. Inmunoprecipitación de proteínas expresadas transitoriamente con epítopos diferentes a T7..... | 65 |
| 4.4.4. Inmunoprecipitación de RNA (RIP)..... | 66 |
| 5. Procedimientos de biología molecular | 67 |
| 5.1. Extracción de RNA y análisis por RT-qPCR..... | 67 |

| | |
|--|-----------|
| 6. Análisis estadístico | 67 |
| 7. Anticuerpos y oligonucleótidos..... | 68 |
| 7.1. Anticuerpos | 68 |
| 7.2. Oligonucleótidos | 70 |
| Resultados | 75 |
| 1. TCERG1 interacciona con la fosfoproteína nucleolar NOLC1 | 77 |
| 1.1. Identificación de la secuencia necesaria para la interacción entre TCERG1 y NOLC1 | 77 |
| 1.2. La interacción entre TCERG1 y NOLC1 está mediada por RNA..... | 82 |
| 2. TCERG1 se asocia a los cuerpos de Cajal | 83 |
| 2.1. NOLC1 interacciona con coilina..... | 83 |
| 2.2. TERC1 se asocia a coilina | 85 |
| 2.3. Efecto de la inhibición de TCERG1 en la localización de coilina. | 88 |
| 2.3.1. Generación de líneas HEK293T silenciadas para TCERG1 mediante CRISPR/Cas9 | 88 |
| 2.3.2. Efecto de la disminución de la expresión de TCERG1 sobre los CBs en células HEK293T..... | 90 |
| 2.3.3. Efecto de la disminución de la expresión de TCERG1 sobre los CBs en células HAP1. | 92 |
| 2.3.4. La ausencia de TCERG1 no afecta a la expresión de coilina..... | 95 |
| 2.3.5. Efecto de la disminución de la expresión de TCERG1 sobre los CBs en células SH-SY5Y..... | 97 |
| 2.4. La disminución de la expresión de TCERG1 no afecta la localización de otros orgánulos nucleares. | 98 |
| 2.4.1. <i>Speckles</i> | 98 |
| 2.4.2. <i>Gems</i> | 98 |
| 2.4.3. HLB | 99 |
| 2.4.4. Nucléolo | 101 |
| 2.5. La reexpresión de TCERG1 rescata el fenotipo de las líneas silenciadas | 105 |

| | |
|--|------------|
| 2.6. Efecto de la sobreexpresión saturante de TCERG1 en la integridad de los CBs..... | 107 |
| 2.7. La sobreexpresión de TCERG1 no afecta la expresión de coilina. ... | 111 |
| 2.8. La sobreexpresión de TCERG1 no afecta la localización de otros orgánulos nucleares | 112 |
| 2.9. Formación de homocomplejos de TCERG1..... | 115 |
| 2.10. El efecto de TCERG1 sobre los CBs no está mediado por RNA | 116 |
| 3. Efecto del silenciamiento de NOLC1 y coilina en la integridad de los CBs | 120 |
| 3.1. El silenciamiento de NOLC1 afecta a los CBs | 120 |
| 3.2. El silenciamiento de NOLC1 y coilina no afecta la integridad de otros orgánulos nucleares | 122 |
| 4. Papel de TCERG1 en la biosíntesis de snRNPs | 124 |
| 4.1. TCERG1 interacciona con distintos componentes relacionados con la biogénesis de los snRNPs..... | 124 |
| 4.2. Efecto de TCERG1 en la formación de snRNPs. | 126 |
| 4.3. La disminución de la expresión de TCERG1 afecta a la síntesis de snRNPs nacientes. | 130 |
| Conclusiones | 133 |
| Discusión..... | 137 |
| 1. Diferencias en la expresión de TCERG1 alteran la integridad de los CBs .. | 140 |
| 2. Papel de TCERG1 en la biosíntesis de los snRNPs..... | 143 |
| Bibliografía | 147 |

Abreviaturas

| | |
|---------------|---|
| BSA | albúmina de suero bovino |
| CB | <i>Cajal bodies, coiled bodies</i> , cuerpos de Cajal. |
| CBC | <i>cap binding complex</i> , complejo de unión al <i>cap</i> |
| CDK | quinasa dependiente de ciclina |
| CG | componente granular |
| CRISPR | repeticiones palindrómicas cortas agrupadas y regularmente espaciadas |
| Ct. | Extremo carboxilo-terminal |
| CTD | dominio carboxiterminal |
| DAPI | 4',6-diamidino-2-fenilindol |
| DFC | componente fibrilar denso |
| DMEM | Dulbecco's Modified Eagle Medium |
| Dox | doxiciclina |
| DSE | elemento distal de la secuencia |
| DTT | ditiotreitól |
| EDTA | ácido etilendiaminotetraacético |
| EGFP | <i>enhanced green fluorescent protein</i> |
| EGTA | ácido etilenglicoltetracético |
| FA | Formamida |
| FBS | suero bovino fetal |
| FC | centro fibrilar |
| FISH | Hibridación fluorescente in situ |
| Fw | <i>Forward</i> |
| GC | componente granular |
| GST | glutación-S-transferasa |

| | |
|---------------|--|
| GTF | factores generales de transcripción |
| He-Ne | helio-neón |
| HEPES | hydroxyethyl-piperazineethane-sulfonic acid |
| HLB | <i>Histone Locus Bodies</i> |
| hnRNP | ribonucleoproteína heterogénea nuclear |
| HRP | <i>horseradish peroxidase</i> |
| hTR | RNA telomerasa |
| IF | inmunofluorescencia |
| IgG | Inmunoglobulina G |
| IMDM | <i>Iscove's Modified Dulbecco's Medium</i> |
| IP | Inmunoprecipitación |
| Kb | kilobase |
| kDa | kiloDalton |
| LEC | complejo pequeño de elongación |
| LLPS | separación de fases líquido-líquido |
| lncRNA | RNA no codificante largo |
| Lsm11 | <i>like-Sm 11, U7 Small Nuclear RNA Associated</i> |
| M | Molar |
| miRNA | microRNA |
| mRNA | RNA mensajero |
| NaAz | Azida sódica |
| NAPs | <i>NOR-associated patches</i> |
| NE | Extracto nuclear |
| NGS | suero de cabra normal, <i>Normal Goat Serum</i> |
| NLS | <i>nuclear localization signal</i> , señal de localización nuclear |
| NOLC1 | Nucleolar and Coiled-Body Phosphoprotein 1 |
| NOR | regiones organizadoras nucleolares |

| | |
|----------------|--|
| NP-40 | Nonidet fosfato-40 |
| Nt. | Extremo amino-terminal |
| OE | Sobre-expresión |
| PBS | Tampón fosfato salino |
| PCR | Reacción en cadena de la polimerasa |
| PFA | Paraformaldehido |
| PI | ioduro de propidio |
| PIPES | <i>piperazine-N,N'-bis(2-ethanesulfonic acid)</i> |
| PML | leucemia promielocítica |
| PMSF | fenilmetil-sulfonil-fluoruro |
| PSE | elemento proximal de la secuencia |
| RIP | Inmunoprecipitación de RNA |
| rRNA | RNA ribosómico |
| RNAPII | RNA polimerasa II |
| RRM | Motivos de reconocimiento de RNA |
| RS | dominio arginina/serina |
| RT-qPCR | PCR en tiempo real mediada por retrotranscripción |
| Rv | <i>Reverse</i> |
| scaRNA | pequeños RNA específicos del cuerpo de Cajal |
| SEC | súper-complejo de elongación |
| SF1 | <i>splicing factor 1</i> |
| SMN | <i>survival motor neuron</i> |
| snoRNA | RNA nucleolar pequeño, <i>small nucleolar RNA</i> |
| snoRNP | ribonucleoproteínas pequeñas nucleolares, <i>small nucleolar ribonucleoprotein</i> |
| snRNA | RNA nuclear pequeño, <i>small nuclear RNA</i> |

| | |
|---------------|---|
| snRNP | pequeñas ribonucleoproteínas nucleares, <i>small nuclear ribonucleoproteins</i> |
| SRSF2 | <i>Serine and Arginine Rich Splicing Factor 2</i> |
| SSC | <i>Saline-Sodium Citrate</i> |
| TCERG1 | <i>transcription elongation regulator 1</i> |
| TMG | 2,2,7-Trimetilguanosina |
| tRNA | RNA de transferencia |
| UV | ultravioleta |
| WB | <i>western blot</i> |
| WCE | <i>whole cellular extract</i> , extracto celular |
| WRAP53 | <i>WD Repeat Containing Antisense to TP53</i> |
| WT | <i>Wild Type</i> /silvestre |

Algunos términos han sido incluidos en su lengua de origen y escritos en cursiva por no tener una clara traducción al castellano o ser ésta poco empleada.

Introducción

1. Arquitectura funcional del núcleo de la célula eucariota

El núcleo de la célula eucariota se caracteriza por contener el material genético de la célula así como las maquinarias moleculares necesarias para su correcta expresión, replicación y reparación. Al contrario de los orgánulos citoplasmáticos, los orgánulos nucleares se caracterizan por la ausencia de una membrana que los separe de los otros componentes del nucleoplasma. A pesar de este hecho, los orgánulos nucleares no están contenidos de manera caótica dentro del volumen nuclear. El núcleo de la célula eucariota es una estructura altamente compartimentada donde el material genético se organiza en cromosomas que ocupan zonas específicas, denominan territorios cromosómicos, dejando espacio nucleoplásmico para un gran número de subestructuras u orgánulos nucleares. Las distintas funciones asociadas al núcleo celular se encuentran espacialmente ordenadas de manera muy definida en el espacio intercromatínico y, por tanto, la disposición y relaciones mutuas a nivel tridimensional de estos orgánulos juegan un papel regulador muy importante. Esta organización del núcleo celular eucariótico, que aún hoy en día es poco entendida, es crucial para que se den lugar un gran número de funciones imprescindibles para la supervivencia de la célula, como son la regulación génica, la reparación del ADN y una correcta división celular.

Los orgánulos nucleares aparecen con una gran variedad de formas y algunos de ellos pueden incluso tener estructuras uniformes mientras otros pueden formar subcompartimentos. Estos orgánulos contribuyen a la regulación de las funciones celulares de manera diferente. Así, los cuerpos de Cajal (CBs) están implicados en la modificación y ensamblaje de los complejos U snRNPs, necesarios para el correcto procesamiento del precursor del RNA mensajero (pre-

mRNA); los *Gems* son estructuras adyacentes o coincidentes con los cuerpos de Cajal caracterizados por la presencia de las proteínas SMN y Gemins 2-7, y relacionados con el ensamblaje de los snRNPs; el nucléolo, implicado en la biogénesis de los ribosomas y la respuesta al estrés celular; el cuerpo PML, implicado en el ensamblaje de la proteína supresora de tumores leucemia promielocítica (PML), además de otras funciones celulares como el secuestro y liberación de proteínas, modificaciones postraduccionales y respuesta al estrés celular; los *paraspeckles*, importantes para el control de la expresión génica mediante la retención de ARN de doble cadena e involucrados en procesos de diferenciación, infección viral o respuesta a estrés; y los *speckles* que como se detallará a continuación, son estructuras nucleares relacionadas con el procesamiento del ARNm (Fig. I1).

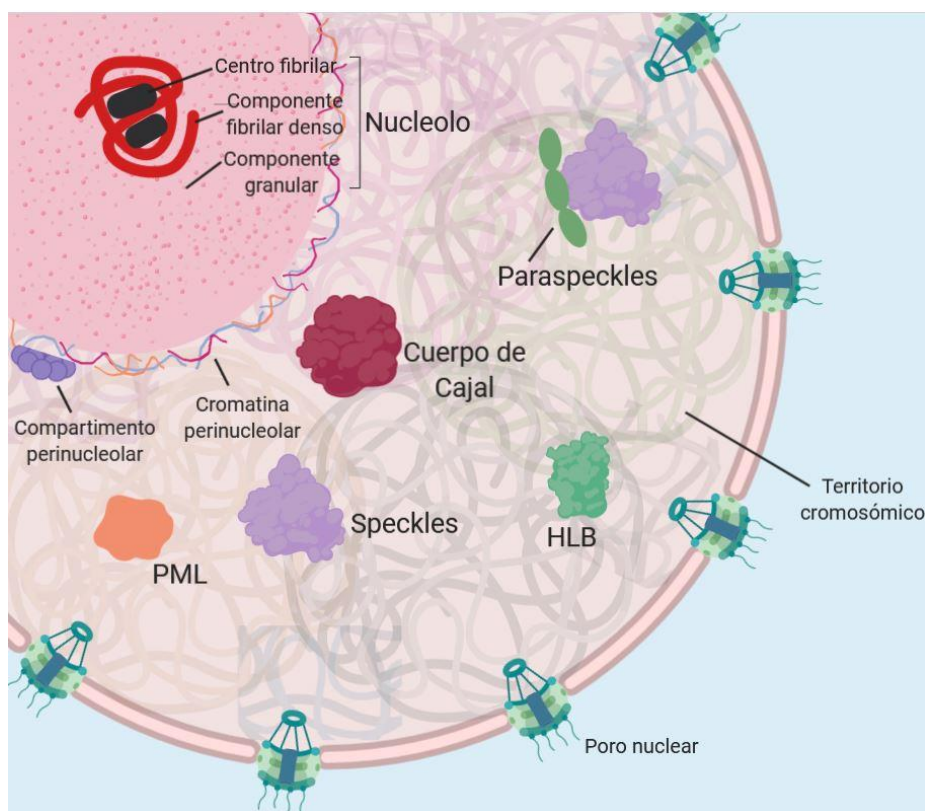


Figura I1. Detalle del núcleo eucariota. La imagen muestra el núcleo interfásico de una célula de mamífero. El núcleo está contenido en una estructura de doble membrana denominada

membrana nuclear, que funciona como una barrera física para separar el contenido nuclear del citoplasma. A lo largo de la membrana nuclear, se encuentran los poros nucleares que controlan el transporte núcleo-citoplasma. La cara interna de la envuelta nuclear proporciona soporte mecánico y participa en la organización de la cromatina. El núcleo contiene la mayoría del material genético de la célula, organizada en cromosomas, que en interfase ocupan los distintos territorios cromosómicos. Las fibras de cromatina pueden contactar entre ellas estableciendo múltiples uniones. El espacio intercromatínico, aunque está muy organizado, es altamente dinámico y en él se encuentran diferentes orgánulos nucleares, como los CBs, HLB, *speckles*, *paraspeckles*, PML o nucléolo. El nucléolo se compone del centro fibrilar, el componente fibrilar denso y el componente granular, que está rodeado de la heterocromatina perinuclear. (Imagen adaptada de (Mao, Zhang, and Spector 2011) realizada en Biorender App).

1.1. Formación de orgánulos nucleares

Los recientes avances en la microscopía de alta resolución y los nuevos conocimientos sobre la importancia de las interacciones entre ARN y proteína y la separación de fases líquido-líquido (LLPS) han permitido entender mejor la estructura y el ensamblaje de los orgánulos nucleares. Se conoce que específicos RNAs son los responsables de concentrar determinadas proteínas para formar orgánulos específicos, como es el caso de NEAT1, un RNA no codificante largo (lncRNA) que lleva a la formación de los *paraspeckles* (Adriaens et al. 2016). Asimismo, los genes de rRNA participan en la formación del nucléolo (Falahati et al. 2016). También hay estudios que implican al RNA en la formación de los CBs. De hecho, se ha descrito que los snRNAs, tanto los recién sintetizados como los maduros, interactúan directamente con distintas proteínas presentes en los CBs haciendo que estas se dirijan a este orgánulo (Machyna et al. 2014). Otros estudios relatan la importancia del mRNA de histonas para la formación del HLB (Heyn et al. 2017). En este sentido, ha sido de especial importancia los estudios realizados con RNAsas para identificar proteínas cuya localización en determinados

orgánulos es dependiente del RNA. Este es el caso de proteínas que se localizan, por ejemplo, en los *speckles* nucleares o en los CBs (Mannen et al. 2016).

Varios estudios han mostrado que las interacciones que ocurren en la formación de orgánulos nucleares pueden ser explicadas por las transiciones de fase líquido-líquido (LLPS) que establecen una variedad de diferentes dominios promotores de agregación en proteínas localizadas en distintos orgánulos nucleares. Estos tipos de interacciones pueden explicar las propiedades líquidas de los orgánulos nucleares, como es su rápida dinámica de unión y ruptura. Se ha visto que en algunas ocasiones la presencia de RNA es determinante para mediar en la formación de distintos orgánulos nucleares a través de las LLPS (Lin et al. 2015; Schwartz et al. 2013). Se han estudiado diferentes combinaciones de elementos que interactúan para mostrar como algunos dominios proteicos solo se someten a LLPS cuando se combinan en concentraciones específicas. De este modo, distintos orgánulos celulares, y entre ellos orgánulos nucleares, podrían compartir componentes, pero cada orgánulo nuclear podría tener diferentes proporciones de proteínas que conducen a LLPS y alterar así el reclutamiento de moléculas secundarias (Banani et al. 2016). Estas proporciones definidas de los elementos proteicos podría explicar las concentraciones variables de los componentes entre los diferentes CBs encontrados en el mismo núcleo (Platani et al. 2000).

1.2. Nucléolo

El nucléolo es la subestructura de mayor tamaño del núcleo, donde se da lugar la biogénesis de los ribosomas, un proceso fundamental para el ensamblaje de la maquinaria de traducción, el ribosoma. Su biogénesis se inicia por la transcripción de los genes RNA ribosomales (rRNA), que se encuentran también

en el nucléolo. En el genoma de las células eucariotas, los genes rRNA presentan generalmente un gran número de copias organizados en series de repeticiones entre diferentes cromosomas en las llamadas regiones organizadoras nucleolares (NORs) (Pelletier, Thomas, and Volarevic 2018). El número de repeticiones de genes rRNA varía enormemente entre los distintos organismos, ya que en humanos y ratones existen unos 200 por genoma haploide, mientras que anfibios y plantas muestran miles de copias. El número de genes ribosómicos varía incluso entre individuos de la misma especie o células de un mismo individuo (Tsekrekou, Stratigi, and Chatzinikolaou 2017; Porokhovnik and Lyapunova 2019). El DNA ribosómico es la parte más transcrita del genoma, produciendo más del 70% de todo el RNA en la célula eucariota (Warner 1999).

El nucléolo es responsable de numerosas funciones y procesos biológicos. Entre estos procesos, se encuentra el mantenimiento de la estabilidad genómica, la reparación del DNA y la recombinación, el control del ciclo celular, el mantenimiento de la inactivación del cromosoma X, la regulación transcripcional y la regulación de estados epigenéticos y al estrés; siendo clave en la organización nuclear (Hutten et al. 2011; Harding, Boiarsky, and Greenberg 2015; Xie et al. 2012; Ochs and Press 1992).

El nucléolo es un orgánulo sin membrana que consta del centro fibrilar (FC), el componente fibrilar denso (DFC) y el componente granular (GC). Se piensa que la transcripción del rRNA se da lugar entre las zonas FC y DFC, mientras que el procesamiento del pre-rRNA ocurre en la región DFC y el ensamblaje pre-ribosomal se da lugar en la región GC (Tiku and Antebi 2018; Boisvert et al. 2007). El tamaño del nucléolo se correlaciona con la síntesis de rRNA, ya que las células en proceso de división, producen grandes cantidades de rRNA y a menudo poseen nucléolos de gran tamaño, mientras que la inhibición de la transcripción de los genes ribosómicos da lugar a una disminución del tamaño nucleolar (Derenzini et al. 1998). De hecho, una alta tasa de transcripción de rRNA y biogénesis de

ribosomas es un rasgo común en muchos tipos de cáncer, usándose incluso el tamaño del nucléolo como un marcador de diagnóstico (Derenzini, Montanaro, and Trere 2009; Hein et al. 2013).

1.2.1. NOLC1

La fosfoproteína nucleolar NOLC1, también conocida como Nopp140, se identificó por primera vez como una proteína de señalización nuclear que actúa como chaperona entre el nucléolo y el citoplasma (Thomas Meier and Blobel 1992). La interacción de NOLC1 con coilin, una proteína presente en los CBs (Andrade et al. 1991), sugiere una posible función de NOLC1 en las conexiones existentes entre el nucléolo y los CBs (Isaac, Yang, and Thomas Meier 1998).

NOLC1 pertenece a una familia de proteínas que se caracterizan por la presencia de un dominio SRP40 en su extremo carboxilo terminal (Yang et al. 2004). Su estructura se caracteriza por la presencia de un dominio central rico en residuos de serina alternadas con secuencias repetidas de lisina, arginina y prolina. La mayoría de las serinas de este dominio central son fosforiladas por la quinasa caseína II, haciendo a NOLC1 una de las proteínas más fosforiladas de la célula con aproximadamente 80 fosfatos por molécula (Thomas Meier and Blobel 1992; Meier 1996). NOLC1 se localiza en el componente fibrilar denso del nucléolo (DFC) y participa en la regulación de la transcripción del RNA ribosómico interactuando con la subunidad grande de la RNA polimerasa I (Chen et al. 1999). Consistente con estos datos, se estableció que NOLC1 interacciona con otros componentes de las maquinarias de transcripción procesamiento de los transcritos del pre-rRNA, como NAP57, diferentes snoRNPs o el factor de transcripción TFII-I (Chen et al. 1999; Yang et al. 2000; Miao et al. 1997; Meier and Blobel 1994). Todos estos datos sugieren la implicación de NOLC1 en los procesos de biogénesis y transporte de los snoRNPs. Además, se ha descrito la participación de NOLC1 en la regulación de la progresión del carcinoma nasofaríngeo al unirse al promotor

del gen TP53 y regular con éste la expresión de MDM2, sugiriendo por tanto un papel de NOLC1 como oncogén en tumorigénesis (Hwang et al. 2009). Consistente con estos datos, NOLC1 está regulado por NF- κ B y CREB, que son reguladores de muchos genes involucrados en ciclo celular, apoptosis y tumorigénesis (Gao et al. 2011).

1.3. Cuerpos de Cajal (CBs, *Cajal Bodies*)

A principios del siglo XX, el neurobiólogo Santiago Ramón y Cajal descubrió unas estructuras presentes en el núcleo neuronal que denominó cuerpos accesorios del nucléolo (Cajal 1903) (Fig. I2). Estos cuerpos fueron redescubiertos y denominados *coiled bodies* durante la década de 1960 para finalmente ser llamados cuerpos de Cajal en honor a su descubridor (Gall 2000). La investigación sobre los CBs se aceleró con el descubrimiento de autoanticuerpos procedentes del suero de un paciente con una enfermedad autoinmune que reaccionaba contra una proteína de aproximadamente 80 kDa que localizaba específicamente en los CBs (Raska et al. 1991). Esta proteína se denominó p80/coilina. El tamaño de los CBs varía desde 0,5 μ m a 1,5 μ m de diámetro y el número de estos orgánulos presentes en la célula varía desde 0 a 15 por núcleo (Pena et al. 2001; Berciano et al. 2007), según el estado metabólico de la célula y el momento del ciclo celular en el que se encuentre (Carmo-Fonseca, Ferreira, and Lamond 1993; Carvalho et al. 1999; Platani et al. 2002). Los CBs son estructuras dinámicas que se mueven, se dividen y se unen dentro del nucleoplasma, intercambiando continuamente componentes con distintos orgánulos del nucleoplasma (Sleeman et al. 2003).

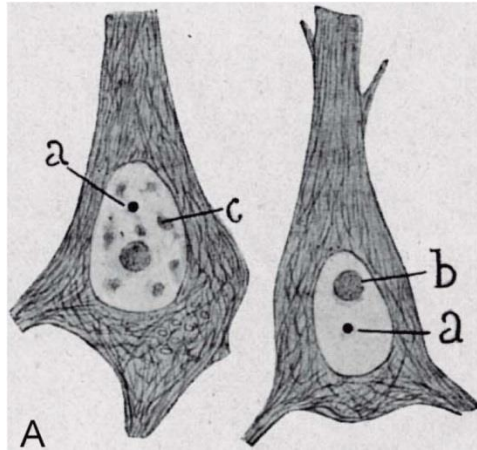


Fig. 12. Cuerpos neuronales. Se muestra el dibujo realizado por Santiago Ramón y Cajal mostrando dos neuronas piramidales de la corteza cerebral humana (Cajal 1903). Se representa el cuerpo accesorio (a), el nucléolo (b) y los grumos hialinos que más tarde se denominaron *speckles* nucleares (c). © de la imagen, *Herederos de Santiago Ramón y Cajal*.

Desde su descubrimiento, la proteína coilina se considera la proteína que define los CBs. El extremo amino y carboxilo de coilin están muy conservados evolutivamente mientras la región central de la proteína y la caja RG parecen ser específicas de vertebrados. El dominio donde se encuentran los primeros 92 aminoácidos de la proteína, parece ser imprescindible para que se dé lugar la autoasociación de coilina, que es a la vez esencial para que se dirija a los CBs (Wu, Murphy, and Gall 1994; Bohmann, Ferreira, and Lamond 1995; Hebert and Matera 2000). La parte central contiene dos señales de localización nuclear y una aparente señal de localización nucleolar que podría explicar su afinidad hacia el nucléolo (Hebert and Matera 2000). En el extremo carboxiterminal se encuentra un dominio Tudor con monometil-lisina, trimetil-lisina y dimetil arginina y además se encuentra la secuencia necesaria para la interacción entre coilin y las proteínas Sm.

Aunque tanto en plantas como en la mosca de la fruta no se detectaron problemas de viabilidad o fertilidad (Collier et al. 2006; Liu et al. 2009), la inhibición de coilina en modelos murinos y de pez cebra mostró la importancia de

esta proteína en fecundidad y embriogénesis (Strzelecka et al. 2010; Walker, Tian, and Matera 2009). La inhibición de coilina en pez cebra es letal y este fenotipo se rescata al inyectar snRNPs maduros, apoyando el modelo de que los CBs promueven el ensamblaje de los snRNPs (Klingauf, Stanek, and Neugebauer 2006; Novotny et al. 2011). Los CBs están asociados a los genes de snRNAs de una forma dependiente a su transcripción, aunque los mecanismos de esta asociación aún no están claros.

Los CBs acumulan numerosos RNAs cortos no codificantes, incluyendo snoRNAs y snRNAs. Además, se ha detectado en estos orgánulos la presencia de otros ARNs cortos, como son los scaRNAs, los cuales son ARNs que guían las modificaciones de los snRNAs. En el caso de los snoRNAs, se ha observado que varios snoRNAs, como es el caso de snoRNA U3, pasan por los CBs antes de llegar al nucléolo, y se ha comprobado que distintos snoRNAs interaccionan directamente con la proteína coilin (Machyna et al. 2014). Estos datos indican que los CBs están involucrados en la biogénesis de snoRNPs (Boulon et al. 2004; Machyna et al. 2014). Varios estudios sugieren que los CBs están involucrados en el procesamiento y modificaciones de los snRNAs, así como en el ensamblaje de los snRNPs funcionales y su reciclaje (Schaffert et al. 2004; Stanek and Neugebauer 2004; Stanek et al. 2008; Stanek et al. 2003; Nestic, Tanackovic, and Kramer 2004; Darzacq et al. 2002; Jady et al. 2003). La importancia de los CBs en la biogénesis de los snRNPs y, dada su relevancia en esta tesis, se discute en profundidad en el apartado 2 de la Introducción.

Todos estos datos sugieren que coilina promueve la biogénesis de las ribonucleoproteínas actuando como chaperona de los pequeños RNAs nucleares no codificantes. Esto podría no ser esencial en células diferenciadas o líneas celulares, pero sí podría ser crítico en embriones en desarrollo. Los datos de inhibición de coilina obtenidos en el pez cebra apoyan esta sugerencia, ya que son incapaces de completar la embriogénesis y presentan defectos en *splicing* y un

reducido número de snRNPs maduros (Strzelecka et al. 2010; Boulon et al. 2004). Destacar que los embriones del pez cebra requieren una alta demanda de snRNPs para que se dé el proceso de *splicing* (Konig et al. 2007). Igualmente sucede en embriones de ratón, donde existe una viabilidad baja en ratones deficientes en coilina y se cree que las pocas células embrionarias que sobreviven se debe a los snRNPs maternos (Machyna, Neugebauer, and Stanek 2015). En otros organismos, como plantas o insectos, las diferencias en los mecanismos de desarrollo podrían determinar si coilina es esencial o no. Por ejemplo, los insectos se caracterizan por un blastodermo sincitial que podría dar una ventaja de supervivencia en el embrión. Si el ensamblaje de los snRNPs es más bajo en los mutantes nulos de *Drosophila melanogaster*, el núcleo del blastodermo puede compartir el déficit de snRNPs entre todas las células de la bástula durante el periodo crítico cuando se necesita un proceso rápido de *splicing*. En este caso, también se baraja la posibilidad de que observando en *D. melanogaster* las diferencias existentes los dominios generalmente conservados de coilina, es posible que los mecanismos implicados en el ensamblaje de los snRNPs puedan no depender tan fuertemente de coilina como ocurre en vertebrados (Deryusheva and Gall 2009)

Se ha sugerido que coilina y los CBs pueden estar implicados en el ensamblaje y el mantenimiento de la telomerasa. Hay estudios en células humanas que localizan la RNA telomerasa (hTR) en CBs durante la fase S del ciclo celular (Jady, Bertrand, and Kiss 2004; Jady et al. 2006; Lukowiak et al. 2001; Zhu et al. 2004; Tomlinson et al. 2006). Este hecho parece ser una situación específica en humanos, ya que no se ha encontrado la hTR en los CBs de ratón (Tomlinson et al. 2010). La telomerasa es retenida en los CBs a través de la proteína WRAP53, un componente de los CBs, uniéndose a la secuencia de la caja CAB del componente RNA de la telomerasa e interaccionando a la vez con la proteína coilina, estableciendo una conexión entre la telomerasa, WRAP53 y coilina (Enwerem et al. 2014; Mahmoudi et al. 2010; Jady, Bertrand, and Kiss 2004; Venteicher et al. 2009;

Tycowski et al. 2009; Richard et al. 2003; Zhong et al. 2011). Y una menor expresión de coilin está asociada con una menor asociación entre la telomerasa y los telómeros y por tanto con una menor actividad de la telomerasa (Venteicher et al. 2009; Stern et al. 2012; Zhong et al. 2012). No obstante, se ha observado que células cancerígenas humanas donde no se expresa coilin no muestran una menor actividad de la telomerasa, ni defectos en la longitud de los telómeros, por lo que se cree que estas células pueden haberse adaptado a la pérdida de coilin y suplen su función sobreexpresando la telomerasa (Stern et al. 2012; Chen et al. 2015). Dado que WRAP53 es una proteína que dirige factores a sitios de rotura de doble cadena en el DNA, la interacción entre coilin y WRAP53 también explicaría que coilin se encuentra en sitios donde hay daño en el ADN y que los CBs son sensibles a la radiación por UV (Bartova et al. 2014; Cioce et al. 2006).

1.3.1. Formación de los CBs

Coilin es un componente esencial de los CBs, ya que una inhibición completa de la proteína da lugar a la desintegración de estos orgánulos en un gran número de organismos diferentes (Collier et al. 2006; Liu et al. 2009; Tucker et al. 2001; Walker, Tian, and Matera 2009). El dominio esencial para la formación de los CBs son los 92 aa del extremo aminoterminal, necesario para la autointeracción de coilina y la formación *de novo* de los CBs (Wu, Murphy, and Gall 1994; Bohmann, Ferreira, and Lamond 1995; Hebert and Matera 2000). El dominio carboxiterminal interacciona con las proteínas Sm que forman parte del complejo snRNP (Toyota et al. 2010; Xu et al. 2005; Hebert et al. 2001). Asimismo, se ha descrito que el nivel de expresión de los snRNPs es un factor importante para la formación de CB (Sleeman, Ajuh, and Lamond 2001).

El proceso de formación de los CBs no es conocido. La coilina está considerada como el factor que enlaza todos los componentes del orgánulo a

través de numerosas interacciones proteína-proteína (Machyna, Heyn, and Neugebauer 2013). Así, el complejo pequeño de elongación o LEC, compuesto por las proteínas ELL, ICE1 e ICE2 en células de mamífero, colocaliza con la coilina en los CBs para, facilitar la síntesis de los snRNAs por la RNAPII (ARN polimerasa II) (Hu et al. 2013; Smith et al. 2011; Polak et al. 2003). Los componentes de LEC interaccionan y colocalizan con la isopeptidasa SUMO, USPL1, que es también un componente esencial de los CBs. Como sucede con la coilina, la disminución de la expresión de USPL1 da lugar a una disgregación de los CBs, reduciendo la transcripción de los snRNAs, y alterando por tanto el *splicing* del mRNA (Hutten et al. 2014a). Más recientemente, se ha determinado que la snRNP 7SK, compuesta por el snRNA 7SK y las proteínas Larp7 y MePCE, se asocia específicamente con LEC para activar la síntesis de los snRNAs y los snoRNAs (Egloff et al. 2017). Se ha observado que la formación de los CBs es dependiente de temperatura, consistente con el modelo LLPS (Carmo-Fonseca, Ferreira, and Lamond 1993).

Se considera a la proteína coilina como el anclaje de los CBs a través de su dimerización y la interacción con otros elementos, como las proteínas Sm presentes en los snRNPs o la proteína SMN responsable del ensamblaje del snRNP (Hebert et al. 2001; Xu et al. 2005). Aunque se ha observado que cualquier componente de los CBs (coilin, proteínas Sm, scaRNAs, SMN) a una localización nuclear específica es suficiente para inducir a la formación de los CBs, es necesaria una serie no secuencial de asociaciones de proteínas para formar la estructura *de novo* (Kaiser, Intine, and Dundr 2008). Consistente con la hipótesis de que la formación de CBs no se da de una manera jerárquica y es necesario establecer multivalentes interacciones con alguna de las moléculas involucradas para la formación de CBs. Aparentemente, los CBs se forman a través de interacciones entre proteína y RNA entre los componentes de estos orgánulos, uniéndose entre ellos directa o indirectamente a través de coilina y SMN, ya que se ha observado

que estas proteínas podrían facilitar la estabilización de las interacciones entre los componentes de los CBs (Tucker et al. 2001).

La fosforilación puede regular la formación de los CBs. Coilina se fosforila en al menos siete residuos y en dos sitios adicionales durante la mitosis (Carmo-Fonseca, Ferreira, and Lamond 1993). El aumento en la cantidad de residuos fosforilados de coilina está relacionado con la disminución de la autoasociación de esta proteína y el desensamblaje de los CBs durante la mitosis (Hebert and Matera 2000). Además, el estado de fosforilación de otras proteínas presentes en los CBs, como SMN (*survival motor neuron*) y WRAP53 (*WD Repeat Containing Antisense To TP53*), influye en su interacción con coilina y la formación de los CBs. La fosforilación puede, asimismo, regular la localización de coilina en el nucleolo (Hebert and Matera 2000; Toyota et al. 2010; Hebert 2013). A pesar del hecho de que coilina no contiene un motivo de unión al RNA canónico, se han encontrado numerosos RNAs que interactúan con coilina y forman parte, por tanto, de los CBs (Machyna et al. 2014). Y de hecho, la fosforilación de la coilina, podría regular al unión al RNA o la asociación con complejos que contengan RNA (Broome et al. 2013). Además de la fosforilación, coilina y otros componentes de los CBs se modifican por las metiltransferasas PRMT5 y PRMT7, que juegan un papel en el control de la formación y actividad de los CBs (Boisvert et al. 2002; Clelland et al. 2009).

1.4. *Gems*

En la década de 1990, se identificó un orgánulo asociado a los CBs denominado *Gem* (Géminis de CBs). Estos cuerpos nucleares fueron identificados por inmunofluorescencia como un orgánulo libre de coilin, pero que contenía las proteínas SMN y Gemins (Liu and Dreyfuss 1996; Liu et al. 1997). SMN se localiza

en el núcleo y en el citoplasma, forma parte junto con gemins de un complejo implicado en la biogénesis de los snRNPs (Charroux et al. 1999), y juega un papel esencial en la formación de los *gems* nucleares (Liu and Dreyfuss 1996; Burlet et al. 1998; Renvoise et al. 2006), por lo que se le ha considerado la proteína que define los *Gems*. Además de su interacción con gemins, SMN se asocia a coilina a través de WRAP53, proteína que es esencial para la localización del complejo SMN en los CBs (Hebert et al. 2001; Mahmoudi et al. 2010). Aunque los *Gems* se encuentran frecuentemente asociados con los CBs, especialmente en células neuronales, y SMN y Gemins colocalizan con coilin y los snRNPs, en otros tipos celulares no siempre sucede así y los *Gems* y los CBs como estructuras separadas dentro del nucleoplasma (Liu and Dreyfuss 1996; Carvalho et al. 1999; Pena et al. 2001).

La pérdida de la expresión de SMN da lugar a una degeneración de motoneuronas y al desarrollo de la atrofia muscular espinal (SMA), una enfermedad que afecta a niños (Lefebvre et al. 1995) (Genabai et al. 2015). Esta enfermedad se produce por deleciones o mutaciones en el gen SMN1, que codifica a la proteína SMN. Debido a una duplicación y una inversión génica, en humanos se encuentra además el gen SMN2 que codifica una forma truncada de la proteína, en la cual falta el exón 7 de SMN (SMN Δ 7) (McAndrew et al. 1997; Lorson et al. 1999). Aunque SMN2 no compensa por completo la pérdida de SMN1, se ha visto que una sobre-expresión de SMN2 reduce la severidad de la enfermedad (Lorson et al. 1999; Le et al. 2005).

1.5. HLB (*histone locus body*)

El HLB es un orgánulo nuclear descubierto por primera vez en *Xenopus laevis* (Wu and Gall 1993). Inicialmente, estos orgánulos fueron identificados como CBs por la presencia de coilin (Frey and Matera 1995). Posteriormente, estudios en

Drosophila melanogaster desvelaron que el HLB es un orgánulo diferente al CB (Liu et al. 2006). En esos estudios, se vió que el snRNP U7 se localizaba únicamente en el HLB y no en los CBs. U7 snRNP se compone del snRNA U7 unido a un anillo de 5 proteínas Sm (B, D3, G, E, F) y dos proteínas LSM (10, 11), exclusivas de esta ribonucleoproteína (Pillai et al. 2003; Pillai et al. 2001). Confirmando la diferencia entre ambos orgánulos, se comprobó en células de mamífero que el HLB era distinto del CB, y se demostró que el ensamblaje de HLB tenía lugar en ausencia de coilin (Liu et al. 2009; Bongiorno-Borbone et al. 2008; Barcaroli et al. 2006; Alm-Kristiansen et al. 2008).

El HLB es un orgánulo nuclear asociado a los *loci* de histonas en la cromatina. Los HLB están compuestos por numerosas proteínas implicadas en la transcripción y procesamiento del pre-mRNA de las histonas, por lo que es un orgánulo esencial en la expresión de los genes de histonas (Salzler et al. 2013; White et al. 2011). Las histonas juegan un papel crucial en el empaquetamiento del DNA a la cromatina, y su expresión está restringida a la fase S del ciclo celular, cuando se da lugar el empaquetamiento del recién replicado DNA (Marzluff and Duronio 2002). Defectos en este proceso dan lugar a inestabilidad genómica y generar distintas patologías como cáncer.

1.6. Paraspeckles

Los orgánulos nucleares denominados *paraspeckles* fueron descubiertos durante un estudio de proteómica de purificación nucleolar. En este estudio, se identificó una proteína teóricamente nucleolar (PSPC1, proteína *paraspeckle* 1), que se acumulaba en gránulos repartidos por el nucleoplasma y que no localizaba con otras proteínas marcadoras de orgánulos nucleares conocidos (Andersen et al. 2002; Fox et al. 2002). Hoy en día, se sabe que los *paraspeckles* contienen unas 40

proteínas diferentes (generalmente proteínas que se unen a RNA) y un lncRNA (NEAT1) (Fox et al. 2002; Chen and Carmichael 2009).

Los *paraspeckles* se encuentran en los espacios intercromatínicos del nucleoplasma y cerca de los *speckles* nucleares. Están implicados en numerosos procesos celulares, incluyendo la retención nuclear de algunos RNAs sometidos a edición (cambio de adenosina por inosina) (Chen and Carmichael 2009) y el control de la transcripción génica a través del secuestro de reguladores transcripcionales (Hirose et al. 2014).

1.7. PML-NB (*PML nuclear bodies*)

PML es un gen implicado en una translocación cromosómica con el gen codificante del receptor α del ácido retinoico que da lugar al desarrollo de la leucemia promielocítica (APL) y es el marcador de los orgánulos conocidos como PML/PML-NB. El hecho de la desaparición de los PML-NB en leucemias y tumores sólidos desvela las propiedades de este orgánulo como supresor de tumores. Además, los PML-NBs están implicados en el control de importantes procesos celulares, incluyendo apoptosis, transcripción, reparación del DNA, senescencia y componente del sistema inmunitario innato (Lallemant-Breitenbach and de The 2010; Bernardi and Pandolfi 2007; Scherer and Stamminger 2016). Estudios recientes han confirmado una función de respuesta al estrés, mostrando que los orgánulos PML se inducen por estrés oxidativo y que la proteína PML promueve la degradación de las proteínas mal plegadas (Sahin et al. 2014; Guo et al. 2014).

1.8. *Speckles* nucleares

Los *speckles* nucleares son unos orgánulos con un tamaño aproximado de 0.3-3 μm , existiendo decenas por núcleo. Después del nucléolo, los *speckles* representan el segundo compartimento nuclear más grande en los núcleos de mamíferos, existiendo en numerosas especies diferentes (Spector and Lamond 2011; Hall et al. 2006). Santiago Ramón y Cajal describió por primera vez estos orgánulos denominándolos “grumos hialinos” (Lafarga et al. 2009) y, posteriormente, los *speckles* se redescubrieron por microscopía electrónica, con el nombre de regiones intercromatínicas (IGCs) (Thiry 1993).

Los *speckles* son estructuras intranucleares que se caracterizan por la presencia de una alta concentración de factores implicados en el *splicing* del pre-mRNA localizados en la región intercromatínica del nucleoplasma (Fig. I3). Tradicionalmente, los *speckles* nucleares se han considerado como sitios de almacenamiento o ensamblaje de los componentes implicados en el *splicing*, pero numerosos estudios han demostrado funciones adicionales en cuanto a la regulación de la expresión génica. De este modo, los *speckles* pueden controlar la dinámica de la transcripción y de los factores de procesamiento del RNA actuando como centros de la activación génica e incluso pueden participar en el procesamiento postranscripcional, así como en el acoplamiento de los primeros pasos de la biogénesis del ARNm (Sanchez-Alvarez, Sanchez-Hernandez, and Sune 2011).

Estudios proteómicos han revelado la composición de los *speckles* nucleares detectándose unas 150 proteínas, muchas de las cuales tienen funciones conocidas en el *splicing* del pre-mRNA, en el metabolismo del RNA, la transcripción y otros rprocesos celulares (Mintz et al. 1999; Saitoh et al. 2004). Uno de estos factores es SRSF2, antes conocido como SC35, un factor de *splicing* esencial que ha venido siendo considerado el marcador por excelencia de estos cuerpos nucleares.

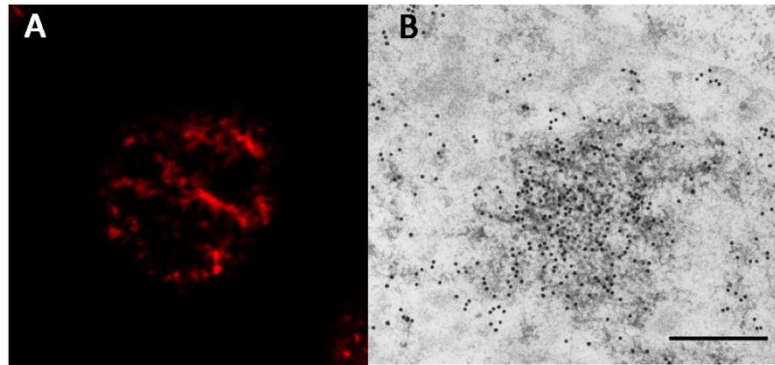


Figura I3. *Speckles* nucleares. (A) Imagen de inmunofluorescencia de células HEK293T empleando anticuerpos específicos frente al factor esencial de *splicing* SRSF2. Se puede observar la distribución de este factor en gránulos nucleoplasmáticos que corresponden a los *speckles* nucleares, con escasa tinción en el nucleoplasma. (B) Imagen obtenida mediante microscopía electrónica en la cual se puede observar con detalle la acumulación de los gránulos intercromatínicos; barra de tamaño: 500nm. (Imagen tomada de (Spector and Lamond 2011)).

1.8.1. Dominios de señalización hacia los *speckles* nucleares

Existen numerosos dominios descritos como responsables de la señalización hacia los *speckles* nucleares. Dentro de estos dominios, se han estudiado con profundidad los dominios arginina/serina (RS) y los dominios RRM (motivos de reconocimiento de RNA). El papel de los dominios RS en la señalización hacia los *speckles* nucleares fue inicialmente descrito en las proteínas de tipo SR Tra y SWAP, ambas reguladoras del proceso de *splicing* en *Drosophila* (Hedley, Amrein, and Maniatis 1995; Li and Bingham 1991). La adición de la secuencia que codifica por el dominio RS en una proteína de localización citoplasmática la conduce a los *speckles* nucleares (Caceres et al. 1997). Otras proteínas de la familia SR, como SRSF2, SRSF3 y SRSF7, presentan el dominio RS, responsable de su localización en los *speckles* nucleares, conteniendo además un único dominio RRM. El dominio RS también es requerido por la quinasa CkrS (Ko, Kelly, and Pines 2001) o el factor de

splicing U2AF⁶⁵ (Gama-Carvalho et al. 1997). Existen otras proteínas SR, como SRSF1 o SRSF10, cuyos dominios RS no son necesarios ni suficientes para la señalización hacia los *speckles* nucleares requiriendo la presencia de, además, al menos un dominio RRM para su localización (Caceres et al. 1997).

Existen otras señales de localización a los *speckles* nucleares. Así, el estudio de las propiedades espaciales de la quinasa DYRK1A demostró la existencia de una cola de histidina responsable de la acumulación de la proteína en *speckles* (Alvarez, Estivill, and de la Luna 2003). Un análisis exhaustivo de las proteínas que contenían esta cola de histidinas, confirmó el papel de este dominio como una nueva señal de localización a los *speckles* nucleares (Salichs et al. 2009). En este estudio, se identificaron 15 proteínas que se localizan en los *speckles* nucleares, llevando a la conclusión que las repeticiones de histidinas son una señal de localización general hacia los *speckles*.

Otras señales de localización a los *speckles* nucleares son “*forkhead-associated*” en NIPP1 (*nuclear regulator of serine/threonine protein phosphatase-1*) (Jagiello et al. 2000), las repeticiones de treoninas y prolinas (dominios TP) de SF3b1/SF3b¹⁵⁵ (Eilbracht and Schmidt-Zachmann 2001) y la región rica en arginina, prolina y serina del factor SRm160 (Wagner, Chiosea, and Nickerson 2003).

Se han propuesto dos modelos diferentes, directo e indirecto, que explican la localización de las proteínas dentro del núcleo. En el modelo directo, las secuencias específicas de señalización median el reclutamiento a través de las interacciones proteína-proteína modulándose, en el caso de los dominios RS, por ciclos de fosforilación y desfosforilación. La localización de las proteínas SR con un único motivo RRM suelen seguir este modelo (Hedley, Amrein, and Maniatis 1995). Como es el caso de la proteína SRSF2, que se localiza en los *speckles* incluso cuando la célula es sometida a tratamientos con RNAsa (Spector, Fu, and Maniatis 1991). Por otro lado, el modelo indirecto contempla que aquellas proteínas que no

contienen una secuencia que dirige hacia los speckles nucleares, se localizan correctamente por la unión a otros componentes que sí tienen la señal de localización pertinente. Por ejemplo, al eliminar los dominios RS o uno de los dos motivos RRM en las proteínas SRSF1 y SRSF5, pueden localizarse en los *speckles* nucleares (Caceres et al. 1997). La localización de estas proteínas en los speckles nucleares probablemente se deba a interacciones con proteínas que sí presenten dominios RS (Hedley, Amrein, and Maniatis 1995).

2. Biogénesis de los snRNPs

Los genes en los organismos eucariotas se transcriben como pre-mRNA que contiene unas secuencias codificantes denominadas exones y unas secuencias no codificantes denominadas intrones. En la década de 1970, se descubrió que los intrones se eliminaban del pre-mRNA mediante un proceso denominado *splicing* para dar lugar al mRNA maduro (Berget, Moore, and Sharp 1977)(Chow et al., 1977). Debido a que la mayoría de los genes humanos contienen numerosos intrones, el *splicing* es un proceso crucial en la expresión génica. Aunque la reacción química del *splicing* es simple, el proceso en sí es complicado y requiere del ensamblaje de un macrocomplejo proteico denominado spliceosoma (Lerner et al. 1980) en varias etapas hasta producir la reacción de transesterificación, corte y ligación de los exones, lo que conlleva un gran gasto de ATP (Berget, Moore, and Sharp 1977).

El spliceosoma se compone de cinco subunidades ribonucleoproteicas (RNPs) asociadas aproximadamente a unas 200 proteínas (Will and Luhrmann 2011; Jurica and Moore 2003). Para distinguirlos de otros RNPs celulares, las subunidades spliceosomales se han denominado ribonucleoproteínas nucleares

pequeñas (*small nuclear RNP*, snRNP) y se componen de cinco ARN pequeños nucleares (snRNAs) ricos en uridinas, U1, U2, U4, U5 y U6 unidas a un anillo formado por 7 proteínas Sm además de otras proteínas específicas. La biogénesis de los snRNPs spliceosomales es un proceso que se da lugar en distintos pasos y en distintos compartimentos celulares, ya que los snRNPs requieren procesos de maduración que tienen lugar en zonas diferentes a los lugares donde desarrollan su función. El ensamblaje de los complejos funcionales y su envío a los destinos finales a menudo están regulados por una serie de complejos intermediarios reguladores ribonucleoproteicos que afectarán a la función de los snRNPs.

2.1. Transcripción y procesamiento de los snRNAs

Los snRNAs son un grupo de transcritos abundantes no codificantes y no poliadenilados que desarrollan su función en el nucleoplasma. En base a su secuencia común, se pueden subdividir en dos clases principales: “Sm snRNAs” y “Sm-like snRNAs”. Aunque las funciones de los snRNAs están bien definidas, la regulación de su expresión todavía no está del todo caracterizada. En metazoos, la transcripción y el procesamiento de los snRNAs están acoplados por un sistema celular que es paralelo, pero distinto de los que generan el RNA mensajero. De hecho, los genes que codifican snRNAs comparten muchas características con los genes que codifican proteínas, incluyendo la posición relativa de elementos que controlan la transcripción y el procesamiento del RNA (Fig. I4). Los snRNAs de tipo Sm se transcriben por la ARN polimerasa II (RNA-Pol II) a través de promotores que contienen en la secuencia elementos distales (DSE) similares a la caja TATA y elementos proximales potenciadores (PSE). Los transcritos de los genes snRNAs no se someten al proceso de *splicing* y no se le añade la cola poliA al extremo 3'. En cambio, la caja 3' característica de los snRNAs se localiza a unas 9-19 bp aguas abajo de la zona codificante, y es requerida para la correcta formación del extremo 3' de los snRNAs (*cap*). Además de los factores generales

de transcripción (factor de inicio de la transcripción IIA (TFIIA), TFIIB, TFIIE y TFIIF), la iniciación de la transcripción de los snRNAs requiere de la unión de un factor pentamérico denominado complejo proteico activador de snRNA (SNAPc) (Henry et al. 1998; Hung and Stumph 2011).

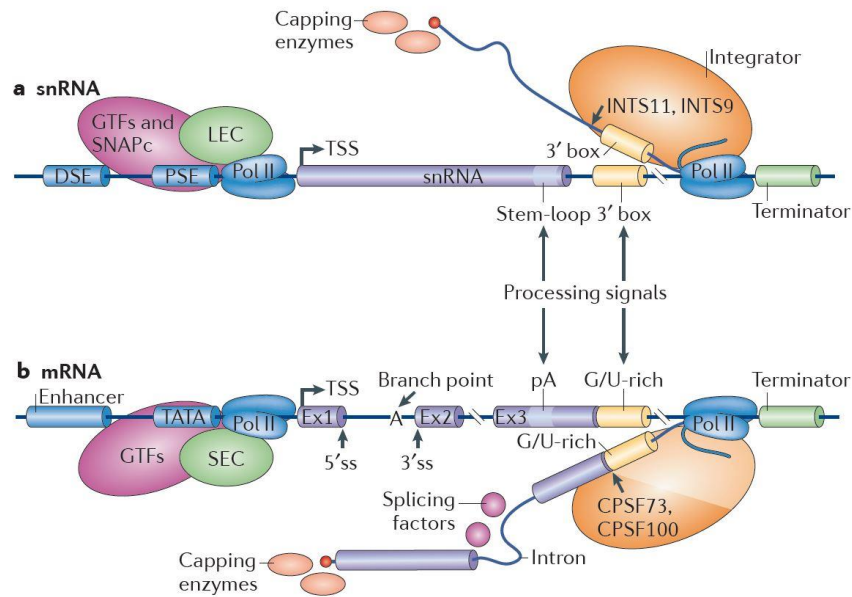


Figura I4. Comparación de la transcripción y procesamiento del snRNA y mRNA. Los genes RNA pequeños nucleares (snRNA) (a) comparten distintas características con los genes que dan lugar a un mRNA que codifica proteínas (b). El elemento de la secuencia distal (DSE) y el elemento de la secuencia proximal (PSE) son equivalentes a los elementos potenciadores (*enhancer*) y la caja TATA, respectivamente, de los genes que dan lugar al mRNA. El promotor del snRNA recluta el complejo pequeño de elongación (LEC), mientras que el promotor del mRNA recluta el súpercomplejo de elongación (SEC). La iniciación de la transcripción del snRNA requiere de factores generales de transcripción así como del complejo proteico activador de snRNA (SNAPc). El complejo *integrator* es requerido para el reconocimiento de las señales de procesamiento aguas abajo, incluyendo la caja 3'. Ex: exón; pA: señal poliadenilización; ss: sitio de *splicing*; TSS: sitio de inicio de la transcripción. (Imagen obtenida de (Matera and Wang 2014).)

Los genes snRNA normalmente presentan múltiples copias, por ejemplo el genoma humano presenta cuatro copias del gen U1 snRNA y 15 copias del gen U2 snRNA (O'Reilly et al. 2013; Lindgren et al. 1985). Una característica fundamental de los genes snRNA es la necesidad del acoplamiento entre los elementos de procesamiento del extremo 3' y el promotor del snRNA, ya que se ha observado que una sustitución del promotor del gen del snRNA por un promotor de una proteína codificante, da lugar a un fallo en la formación del extremo 3' del RNA (Hernandez and Weiner 1986).

Los genes snRNA transcritos por la ARN polimerasa III (7SK y U6) también tienen los elementos DSE y PSE, pero además poseen una caja TATA en la posición -25. Se ha observado que colocando una caja TATA en la posición -25 en los genes snRNA que se transcriben por la ARN polimerasa II, estos pasan a transcribirse por la Pol III (White 2011). Al igual que ocurre con el RNA transcrito por la RNA polimerasa II, la adición de un residuo al extremo 5' (*capping*) del snRNA, así como el corte de su extremo 3', se da cotranscripcionalmente. Para la maduración del extremo 3' del snRNA, se requiere de un complejo que está formado por distintas subunidades denominado complejo *integrator* (Baillat et al. 2005; Chen et al. 2012), que reconoce la señal de procesamiento caja 3' y procesa el transcrito que va saliendo de la polimerasa. Las subunidades 11 y 9 del complejo *integrator* tienen secuencias similares a los factores CPSF73 y CPSF100, respectivamente, que se encargan de la poliadenilación del mRNA (Baillat et al. 2005). Sin embargo, más allá de estas dos subunidades, el complejo *integrator* tiene pocas similitudes con la maquinaria que lleva a cabo el procesamiento del mRNA (Chen et al. 2012). El heterodímero formado por la quinasa dependiente de ciclina 8 (cdk8) y ciclina C, muestra una actividad de procesamiento del extremo 3' del snRNA y está asociado al complejo *integrator* (Baillat et al. 2005).

2.2. Control de calidad y exporte de los snRNAs

La función de los snRNPs se da lugar principalmente en el núcleo. Sin embargo, en la mayoría de las especies, los recién sintetizados snRNAs se exportan en primer lugar al citoplasma donde se someten a diferentes pasos de maduración antes de ser reimportados de nuevo al núcleo. Se han encontrado algunas excepciones, como en levaduras o tripanosomátidos, en los cuales se ha visto que el ensamblaje de los snRNPs se restringe al núcleo (Boon et al. 2007; Murphy, Olson, and Siliciano 2004; Tkacz et al. 2007; Jae et al. 2011; Palfi et al. 2009). La mayoría de los distintos tipos de RNA, incluyendo RNA ribosomal, tRNA, mRNA y microRNA (miRNA) son exportados al citoplasma después de la transcripción y su procesamiento. Típicamente, los ARNs contienen secuencias y/o estructuras específicas que son determinantes para la unión directa o indirecta del receptor apropiado para el transporte, como ocurre para el tRNA y rRNA (Cullen 2003). Debido a que los mRNAs y los snRNAs se transcriben por la ARN polimerasa II, comparten la estructura característica del extremo 5' (*cap*) por lo que se plantea el problema de cómo la maquinaria de exporte diferencia entre estos dos tipos de RNAs. Un serie de estudios sugieren que esta distinción se da lugar por el tamaño del RNA y su asociación con las proteínas C1 y C2 de los complejos ribonucleoproteicos heterogéneos nucleares (hnRNP C1-C2) (Ohno et al. 2002; Fuke and Ohno 2008; McCloskey et al. 2012; Masuyama et al. 2004). Los transcritos por la RNAPII mayores de aproximadamente 250 nucleótidos se unen a los tetrámeros de hnRNP C1-C2 y el exporte se produce a través del factor nuclear de exporte de ARN TAP/NXF1 (McCloskey et al. 2012). Los transcritos más cortos son exportados por un conjunto distinto de factores que incluyen CBP80 y CBP20, que se unen al extremo 5' metilado (*cap*) y forman el complejo de unión al *cap* (CBC) y los factores PHAX y ARS2 (Izaurralde et al. 1994; Ohno et al. 2000; Hallais et al. 2013). Estas proteínas forman una unión entre el *cap* del extremo 5' y el receptor de

exporte CRM1 que interacciona con las proteínas del poro nuclear (Fornerod et al. 1997) (Fig. I5).

Varios estudios sugieren que los precursores de los transcritos de snRNAs (pre-snRNAs) transitan por los CBs antes de salir del núcleo. Por ejemplo, se han encontrado los precursores de los snRNPs en los CBs de mamíferos mediante técnicas de hibridación *in situ* (Smith and Lawrence 2000), así como la acumulación en estos orgánulos de los factores PHAX y CRM1 (Boulon et al. 2004). Además, otros estudios de microinyección en ovocitos de *Xenopus laevis* indican también que los pre-snRNAs se acumulan temporalmente en los CBs, localización que va disminuyendo conforme los snRNAs se van exportando al citoplasma (Suzuki, Izumi, and Ohno 2010; Frey and Matera 2001).

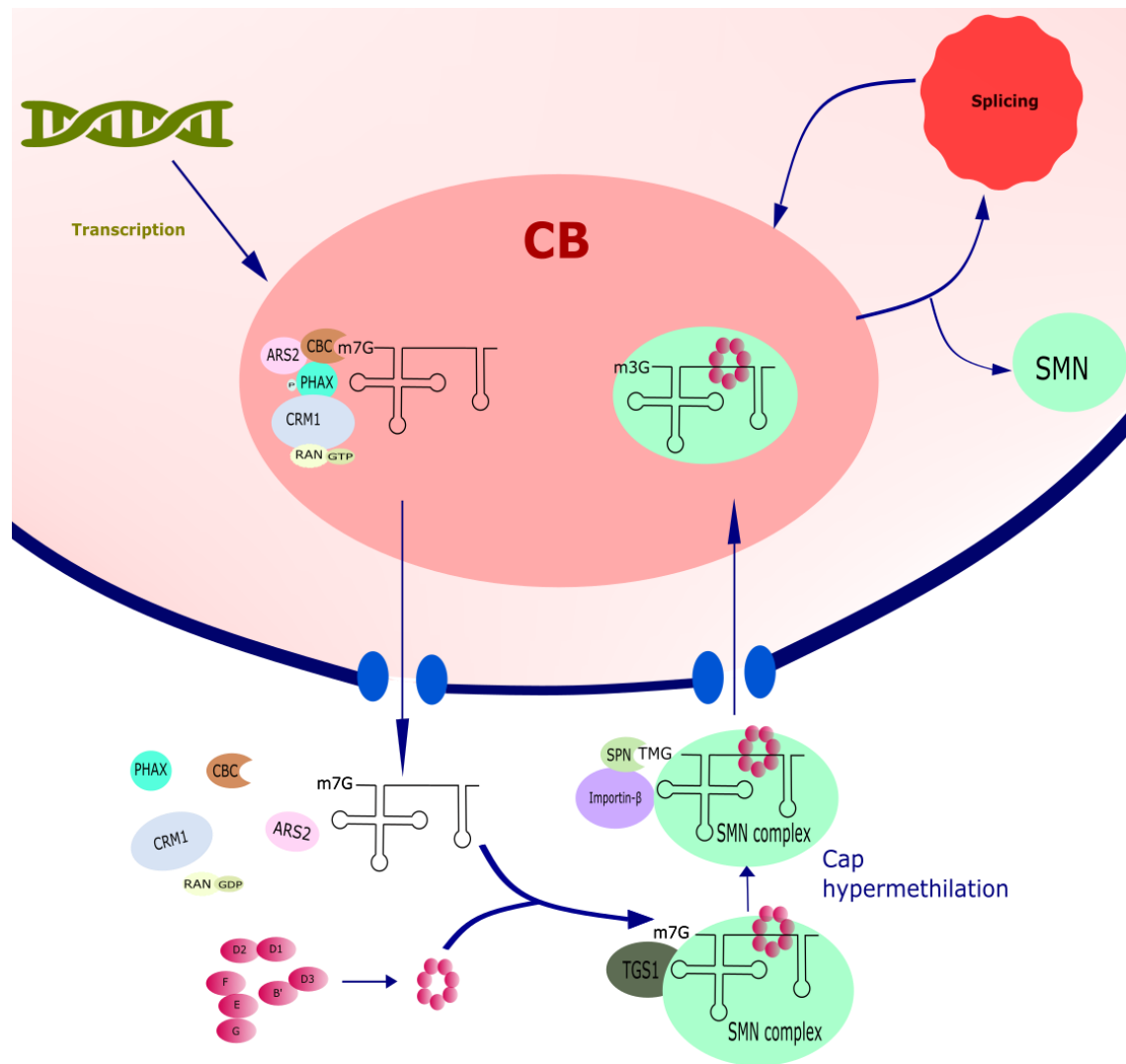


Figura I5. La maduración de los snRNAs requiere distintos pasos citoplasmáticos y nucleares. El complejo de pre-exporte del snRNA consiste en el complejo heterodimérico de unión al *cap* (CBC), la proteína ARS2, la forma hiperfosforilada del adaptador PHAX y un conjunto de proteínas del complejo *integrator*. Este complejo es remodelado en el nucleoplasma, con la separación de este complejo *integrator* y la adición de las proteínas CRM1 y RAN GTPasa. La remodelación incluye un paso por los CBs que se cree funcionan como control de calidad del snRNPs. Una vez que este complejo llega al citoplasma, los factores de exporte se disocian del pre-snRNA antes del ensamblaje de las proteínas Sm. A continuación la proteína TGS1 hipermetila el extremo 5' formando el *cap* 2,2,7-trimetilguanosina (TMG) desencadenando en ensamblaje del complejo de importe nuclear, el cual incluye Snuportin (SPN) y el receptor de importe importina-β, que interaccionan con el complejo SMN. Al entrar de nuevo en el núcleo, el complejo Sm-snRNP pasa por los CBs para que se den lugar unos pasos de maduración. Por último, el complejo SMN se

disocia de los snRNPs antes de ir a los *speckles* nucleares donde se encuentran los factores implicados en el proceso de *splicing*.

2.3. Ensamblaje citoplasmático de los snRNPs

Después de la llegada del pre-snRNA al citoplasma, se produce la disociación del complejo de exporte desencadenado por la desfosforilación de PHAX (Kitao et al. 2008). El complejo de proteínas SMN, el cual incluye la proteína SMN más un conjunto de proteínas asociadas que se denominan gemins, regula la fase citoplasmática del proceso de biosíntesis de snRNPs (Meister, Buhler, et al. 2001; Pellizzoni, Yong, and Dreyfuss 2002; Massenet et al. 2002; Narayanan et al. 2002; Mouaikel et al. 2003). Este complejo proteico de SMN recluta a los recién exportados snRNAs y hace que se unan un conjunto de 7 proteínas Sm a sitios específicos del RNA. Estas proteínas llegan al complejo SMN a través del complejo PRMT5 (Meister, Eggert, et al. 2001). La unión del anillo proteico Sm no solo estabiliza el snRNA, sino que además es requerido para el reimporte del complejo riboproteico de nuevo al núcleo. Las proteínas Sm no se unen al snRNA como un anillo preformado, sino que estas proteínas forman complejos heterodiméricos (SmD1-SmD2 y SmB-SmD3) y heterotriméricos (SmE-SmF-SmG). La proteína *gemin2*, que se encuentra en el complejo SMN, se une directamente a 5 proteínas Sm (SmE-SmF-SmG y SmD1-SmD2) formando una estructura en forma de herradura que se une al RNA, estabilizado por la proteína SMN a través de su dominio Tudor (Grimm et al. 2013). Posteriormente se une el complejo SmB-SmD3 para cerrar el anillo Sm (Zhang et al. 2011).

2.4. Importe nuclear de los snRNPs

La formación del anillo Sm protege y estabiliza el snRNA e inicia una serie de pasos de procesamiento del RNA que termina en el importe al núcleo por el complejo SMN (Fig. I5). Como parte de sus funciones, el complejo SMN recluta la trimetilguanosina sintasa 1 (TGS1), una RNA metiltransferasa que modifica el extremo 5' del snRNA para formar la estructura 2,2,7-trimetilguanosina (TMG) (Mouaikel et al. 2003). Este extremo TMG funciona como una señal de localización nuclear (Fischer and Luhrmann 1990). Antes del reimporte al núcleo, el extremo TMG desencadena la disociación de TGS1 y seguidamente se une Snuportin, el adaptador específico de importe de snRNPs, ya que interacciona directamente con el receptor importin- β , que junto al complejo SMN promueve la entrada del complejo de nuevo al núcleo (Narayanan et al. 2004; Palacios et al. 1997).

El complejo SMN se disocia una vez que el snRNP entra en el núcleo, ya que no se ha encontrado esta proteína en la copurificación de los snRNPs maduros (Neubauer et al. 1998; Trinkle-Mulcahy et al. 2008; Herold et al. 2009). En la mayoría de los distintos tipos celulares, la fracción nuclear de SMN colocaliza principalmente con los CBs, sin embargo y como se ha mencionado previamente, SMN se localiza además en los *Gems* (Matera and Shpargel 2006). Los snRNPs recién importados al núcleo celular se acumulan temporalmente en los CBs antes de almacenarse en los *speckles* nucleares. En los CBs se dan lugar los últimos pasos de maduración de los snRNPs, incluyendo las modificaciones covalentes llevadas a cabo por RNAs no codificantes que se encuentran en estos orgánulos, así como la unión proteínas específicas a los snRNPs a (Jady et al. 2003; Nesic, Tanackovic, and Kramer 2004; Schaffert et al. 2004). En los CBs, no solo se da lugar el ensamblaje *de novo*, sino que se ha encontrado que también ocurre el reciclaje *postsplicing* de los complejos U4-U6 di-snRNP y U4-U6-U5 tri-snRNP (Novotny et al. 2011; Stanek and Neugebauer 2004; Stanek et al. 2003).

2.5. U6 snRNP

El snRNA U6 tiene muchas diferencias con el resto de snRNAs. Es el único snRNA que se transcribe por la RNA polimerasa III y, como resultado, no recibe el *cap* 7-metil guanosina sino que se somete a una metilación en su extremo 5' (Singh and Reddy 1989). El recién sintetizado snRNA U6 no abandona el núcleo de la célula durante su maduración pero se une a la proteína La, a través de su extremo 3', estabilizando el nuevo transcrito. Las modificaciones que se producen en el U6 snRNA están guiadas por los snoRNPs que se encuentran en el nucléolo (Ganot et al. 1999; Tycowski et al. 1998). Consistente con estos datos, se ha observado que el snRNA U6 se localiza temporalmente en el nucléolo (Gerbi, Borovjagin, and Lange 2003). La proteína La es reemplazada por siete proteínas LSm (2-8), que forman un anillo similar a las proteínas Sm. Las proteínas LSm pueden ser transportadas al núcleo como un complejo preformado, lo que respalda aún más el modelo nuclear de maduración del U6 snRNP (Zaric et al. 2005). SART3 es una proteína que contiene dos motivos de reconocimiento a RNA, que se une directamente al U6 snRNA e interacciona a través de un dominio en su extremo carboxiterminal con las proteínas LSm. Se ha observado que, además, interacciona a través de un dominio HAT que se encuentra en el extremo aminoterminal con la proteína Prp3, proteína específica del complejo U4/U6 (Bell et al. 2002; Medenbach et al. 2004; Rader and Guthrie 2002). Estas interacciones sugieren que SART3 podría proporcionar el andamiaje para la formación del snRNP U4/U6.

2.6. Reciclaje de los snRNPs

Durante el *splicing*, los snRNPs sufren una serie de reestructuraciones, en especial los complejos U4/U6 snRNP y U4/U6·U5 snRNP. Los snRNAs U5 y U6

juegan un papel importante en la catálisis del *splicing*, mientras que el U4 snRNA se disocia del spliceosoma antes de su activación (Jurica et al. 2002; Makarova et al. 2004). Por tanto, el proceso de *splicing* disocia el complejo U4/U6 snRNP, por lo que U4 yU6 tienen que ser reensamblados de nuevo para formar el complejo snRNP y poder participar en la siguiente ronda de *splicing*. Se ha podido observar que SART3 y los snRNPs no se localizan en los CBs al inhibir la transcripción y el *splicing*, sugiriendo un tránsito a los CBs para el re-ensamblaje de los snRNPs (Stanek et al. 2003).

3. Proteínas con dominios WW y FF

Existe una familia de proteínas que se caracterizan por la presencia de dominios WW y FF y que han sido relacionados con los procesos de transcripción y *splicing*, ya que establecen conexiones físicas con componentes de ambas maquinarias. La mayoría de estas interacciones se dan lugar a través de los dominios WW y FF, que contienen repeticiones en tándem de secuencias específicas.

3.1. Características de los dominios WW y FF

3.1.1. Dominios WW

Los dominios WW son motivos proteicos pequeños compuestos aproximadamente de 40 aminoácidos que median interacciones de tipo proteína-proteína a través del reconocimiento de secuencias peptídicas ricas en prolina y sitios de serina/treonina fosforilados. Se caracterizan por la presencia de dos residuos de triptófano altamente conservados situados entre los aminoácidos 20 y

22. Su estructura se pliega de manera estable dando lugar a una triple hélice antiparalela con estructura de lámina β en ausencia de ligando (Fig. I6) (Macias, Wiesner, and Sudol 2002). Un mismo dominio puede unirse a varias proteínas diferentes, ya que son capaces de establecer gran variedad de interacciones, lo que los convierte en grandes soportes de conexión entre diferentes procesos celulares. Estos dominios se encuentran presentes en una gran variedad de proteínas envueltas en procesos tales como la transcripción y el procesamiento del RNA, la estabilidad del citoesqueleto, la estabilidad de distintas proteínas y la localización proteica.

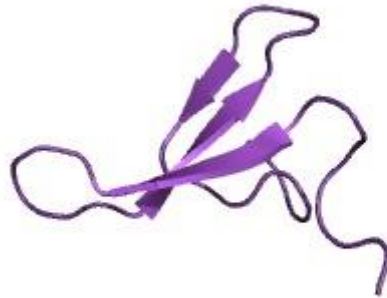


Figura I6. Dominio WW. La imagen representa la estructura secundaria de un dominio WW. Imagen tomada de (Macias, Wiesner, and Sudol 2002).

3.1.2. Dominios FF

Los dominios FF son motivos estructurales de aproximadamente 50-60 aminoácidos, caracterizados por contener dos residuos de fenilalanina altamente conservados situados cerca de sus extremos amino y carboxiloterminales. Se caracterizan por establecer interacciones proteína-proteína en ligandos con secuencias preferentemente fosforiladas. Estos dominios forman una estructura compuesta por tres hélices α , dispuestas ortogonalmente y una hélice 3_{10} entre la

segunda y la tercera hélice α (Fig. I7). Aunque se han definido tradicionalmente como dominios de interacción con las formas fosforiladas del CTD de la RNAPII (Allen et al. 2002; Morris and Greenleaf 2000), los dominios FF son capaces de interactuar con múltiples factores tanto de transcripción como de *splicing* (Gasch et al. 2006; Smith, Kulkarni, and Pawson 2004).

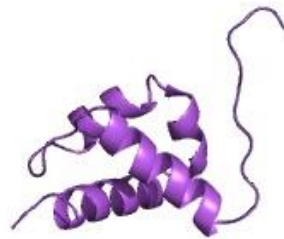


Figura I7. Dominio FF. La imagen representa la estructura de un dominio FF de la proteína PRPF40A. Imagen tomada de (Allen et al. 2002).

Además de las proteínas implicadas en procesos del metabolismo del RNA, existen otras proteínas que presentan estos dominios altamente conservados, citoplasmáticas: la familia de proteínas p190 de las Rho/GTPasas, de localización citoplasmática e implicadas en la regulación de las vías de señalización del citoesqueleto (Bedford and Leder 1999). Las proteínas de la familia RhoGAP no presentan dominios WW que acompañen a los dominios FF, presentando un dominio GTPasa seguido de cuatro dominios FF dispuestos en tándem, y un dominio catalítico RhoGAP en la parte carboxiloterminial (Burbelo et al. 1995).

3.2. PRPF40A

La proteína PRPF40A, también conocida como FBP11 o HYPA, es una proteína de 957 aminoácidos cuya estructura contiene dos dominios WW en su región amino-terminal y seis dominios FF en su extremo carboxilo-terminal (Fig. I8). PRPF40A ha sido localizada en los *speckles* nucleares y se ha publicado que inhibe a la proteína NWASP, implicada en la reorganización del citoesqueleto (Mizutani, Suetsugu, and Takenawa 2004). Además, se ha observado una unión con los factores de *splicing* SF1, U2AF⁶⁵, proteínas de la partícula U2 snRNP, algunos hnRNPs y factores remodeladores de la cromatina, entre otros. Por tanto, PRPF40A se relaciona con el procesamiento del mRNA (Lin, Lu, and Tarn 2004). Además, una caracterización estructural de los dominios FF revelaron que PRPF40A podría reconocer la secuencia fosforilada del CTD de la RNAPII (Allen et al. 2002)

3.3. PRPF40B

La proteína PRPF40B, también conocida como HYPC, es una proteína de 871 aminoácidos que contiene dos dominios WW en su extremo aminoterminal y seis dominios FF en su extremo carboxiloterminal (Fig. I8). PRPF40B colocaliza con SRSF2 en los *speckles* nucleares, siendo el dominio FF4 suficiente para la localización de esta proteína en esta región nuclear. Asimismo, se ha observado que PRPF40B se asocia a los factores de *splicing* SF1 y U2AF⁶⁵ tanto *in vitro* como *in vivo* de una forma dependiente de la integridad del RNA. También se ha observado que PRPF40B es capaz de modular el *splicing* alternativo de genes implicados en apoptosis (Becerra et al. 2015). PRPF40B se ha relacionado con el síndrome mielodisplásico (MDS), un conjunto de desórdenes hematológicos con

una producción desregulada de glóbulos rojos y que muestra una predisposición a la leucemia mieloide aguda (AML) (Yoshida et al. 2011).

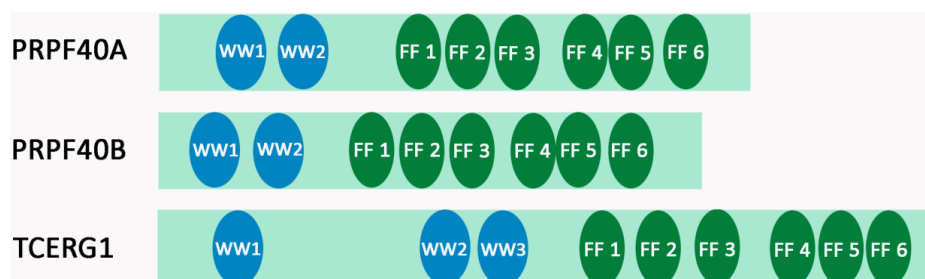


Figura I8. Estructura de las proteínas PRPF40A, PRPF40B y TCERG1. Representación esquemática de las estructuras de las proteínas que se caracterizan por la presencia de dominios WW y FF. PRPF40A y PRPF40B contienen dos dominios WW y seis dominios FF. TCERG1 contiene tres dominios WW y seis dominios FF.

3.4. TCERG1, factor regulador de la elongación transcripcional (*Transcription Elongation Regulator 1*)

TCERG1, previamente conocido como CA150, es una proteína de 1098 aminoácidos con tres dominios WW en su extremo aminoterminal y seis dominios FF en su extremo carboxiloterminal (Fig. I8). Fue aislado por primera vez de extractos nucleares de células HeLa, identificándose como una proteína capaz de regular la transcripción del virus de la inmunodeficiencia humana (HIV-1) mediada por el *trans*-activador viral Tat (Suñé et al. 1997).

En su parte aminoterminal, además de los tres dominios WW, TCERG1 contiene una secuencia rica en residuos de prolina (P), una secuencia rica en residuos de alanina y glutamina (QA) y una secuencia rica en residuos de serina, treonina y prolina (STP) (Fig. I9). El dominio WW1 pertenece al denominado

grupo fosfoserina/fosfotreonina (poS/poT), y se ha descrito que interacciona con el factor de *splicing* SF1 (Goldstrohm et al. 2001; Goldstrohm, Greenleaf, and Garcia-Blanco 2001). El dominio WW2 se engloba dentro del grupo denominado L o poli-P, e interviene en el reconocimiento en factores de *splicing* ricos en prolina, como SF1, U2 snRNP y U2AF (Goldstrohm et al. 2001; Lin, Lu, and Tarn 2004). En cuanto al dominio WW3, se ha descrito que para su correcto plegamiento es necesaria la región α -hélice situada delante del dominio (Fidan et al. 2011) y se ha observado que existe una interacción de este dominio con el factor de *splicing* U2AF⁶⁵ (Sánchez-Álvarez et al. 2006). La secuencia QA está formada por 38 repeticiones de glutamina y alanina (QAQV(QA)₂₉QVQAQVQAQV(QA)₂) (Suñé et al. 1997), siendo susceptible de sufrir expansiones del hexámero de nucleótidos CAGGCC. De hecho, se ha descrito una correlación entre la longitud y la secuencia de las repeticiones de QAs de TCERG1 y la edad a la que se empiezan a manifestar los síntomas de la enfermedad de Huntington (Holbert et al. 2001).

En la región carboxilterminal y además de los seis dominios FF, encontramos una secuencia de cremallera de leucinas (LZ, *leucine zipper*) que solapa de forma parcial con el dominio FF6, una secuencia de poliprolinas semejante a la descrita anteriormente y una secuencia reguladora tipo PEST (secuencia peptídica rica en residuos de prolina (P), ácido glutámico (E), serina (S), y treonina (T)) (Fig. I9). Los dominios FF son los responsables de la localización de TCERG1 en los *speckles* nucleares. En los estudios iniciales, se observó que la ausencia de los tres últimos dominios FF (FF4 al FF6) generaba una proteína truncada que mostraba un patrón difuso por todo el nucleoplasma (Sánchez-Álvarez et al. 2006). Análisis posteriores determinaron que los dominios FF4 y FF5 eran los responsables de la acumulación de TCERG1 en la periferia de los *speckles* nucleares. La inclusión de los dominios FF4 y FF5 a una proteína heteróloga la localizaba específicamente a esta región nuclear (Sánchez-Hernández et al. 2012).



Figura I9. Estructura de TCERG1. Representación esquemática de la estructura de TCERG1. PP: secuencia rica en prolinas; WW: dominios con dos residuos altamente conservados de triptófano; QA: secuencia rica en residuos de glutamina y alanina; STP: región rica en residuos de serina, treonina y prolina; NLS: señal de localización nuclear; FF: dominio con dos fenilalaninas altamente conservadas; LZ: secuencia de cremallera de leucinas; PP-PEST, secuencia peptídica rica en residuos de prolina, ácido glutámico, serina y treonina.

3.4.1. Implicación de TCERG1 en el proceso de transcripción y *splicing*

TCERG1 coimmunoprecipita en extractos nucleares con la RNAPII e interacciona *in vitro* con su CTD hiperfosforilado a través de los dominios FF de la proteína (Suñé et al. 1997; Goldstrohm et al. 2001). Además, TCERG1 se asocia a componentes del complejo PTEFb (ciclina T1 y quinasa CDK9) y a los factores TFIIH y Tat-SF1 (*Tat specific factor 1*) (Smith, Kulkarni, and Pawson 2004; Sánchez-Álvarez et al. 2006), y se localiza a lo largo de las regiones codificantes de genes (Montes, Cloutier, et al. 2012). Todos estos datos apuntan a un papel de TCERG1 durante la elongación transcripcional del mRNA (Suñé et al. 1997)(Suñé and Garcia-Blanco 1999). Apoyando esta hipótesis, TCERG1 afecta la capacidad transcripcional del factor CEBPA (*CCAAT/enhancer-binding protein alpha*) (McFie et al., 2006) y del factor implicado en determinación (*fate*) celular DACH1 (Dachshund Homolog 1) (Zhou et al., 2010).

TCERG1 también interacciona con numerosos factores de la maquinaria de *splicing*. Así, se ha descrito la asociación de los dominios WW con SF1, U2AF⁶⁵ y proteínas del U2 snRNP (Goldstrohm et al. 2001; Sánchez-Álvarez et al. 2006) y la asociación de los dominios FF con Tat-SF1 y proteínas SR (Jablonski et al. 2010).

Funcionalmente, TCERG1 modula el *splicing* alternativo de varios minigenes reporteros (Lin et al, 2004; Cheng et al., 2007; Pearson et al., 2008; Sánchez-Álvarez et al., 2010; Montes et al., 2012) y de numerosos genes endógenos (Pearson et al., 2008; Muñoz-Cobo et al., 2017). En un importante trabajo del laboratorio, se demostró que TCERG1 regula el *splicing* alternativo del gen BCL2L1 modulando la procesividad de la RNAPII (Montes et al., 2012) de una forma dependiente de los dominios FF4 y FF5 (Sánchez-Hernández et al. 2012).

Basado en estos datos, se ha propuesto que TCERG1 acopla la RNAPII elongante con los complejos del spliceosoma para regular eventos cotranscripcionales de *splicing* (Goldstrohm et al. 2001; Montes, Becerra, et al. 2012).

3.4.1. Localización de TCERG1 en los *speckles* nucleares

Consistente con el papel en el procesamiento del pre-mRNA, TCERG1 colocaliza con los *speckles* nucleares. Al analizar la distribución espacial de TCERG1 dentro de estos dominios, se observó que esta proteína se distribuye a lo largo de su periferia de los *speckles*, coincidiendo con los sitios activos de transcripción de los mismos (Fig. I10-A) (Sánchez-Álvarez et al., 2006). Como se ha mencionado, los dominios FF4 y FF5, que comprenden los aminoácidos 878 a 1021, son esenciales para su acumulación en los *speckles* nucleares (Sánchez-Hernández et al. 2012). En este mismo estudio se comprobó que el cambio de los residuos altamente conservados de fenilalanina en posiciones 903 y 961 por alaninas en esos dominios generan una proteína mutante (TCERG1(F903A/F961A)) con una distribución difusa en el nucleoplasma celular, sin acumulación evidente en los *speckles* nucleares (Fig. I10-B) (Sánchez-Hernández et al. 2012).

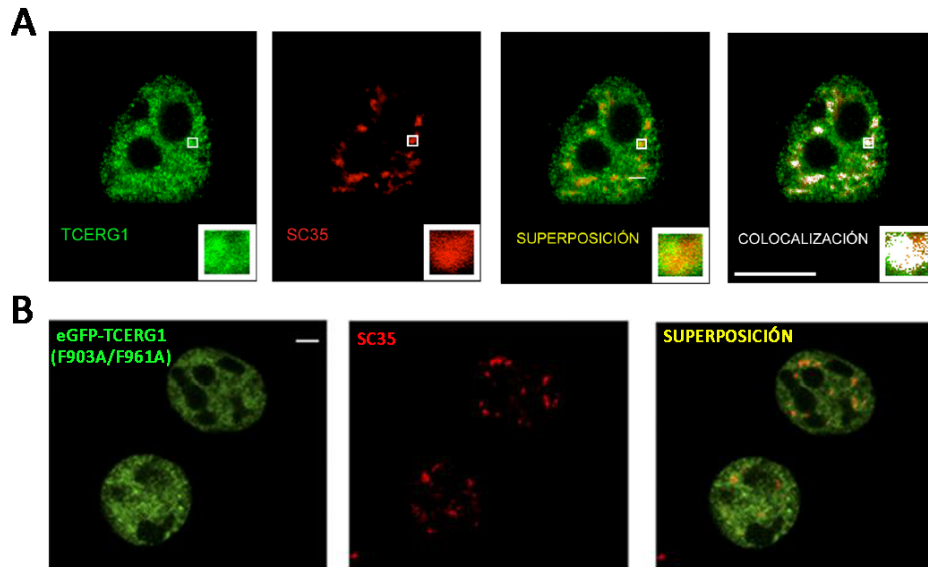


Figura I10. Análisis mediante microscopía confocal de la localización subnuclear de TCERG1 y TCERG1(F903A/F961A). A. Colocalización de TCERG1 en los *speckles* nucleares. Imágenes de TCERG1 (verde), SRSF2 (SC35) (rojo) y la superposición y colocalización de ambas. Imagen obtenida de (Sánchez-Álvarez et al. 2006) B. La proteína mutada TCERG1(F903A/F961A) tiene una distribución difusa por el nucleoplasma. Se muestran imágenes de TCERG1 (verde), SRSF2 (SC35) (rojo) y la superposición de ambas Imagen obtenida de (Sánchez-Hernández et al. 2012).

3.4.2. Interacción de TCERG1 con NOLC1

Previo al inicio de esta tesis doctoral se realizaron ensayos de *pull down* con los dominios FF4/FF5 WT y los mismos motivos conteniendo las mutaciones en los residuos de las fenilalaninas 903 y 961 con el fin de identificar proteínas que pudiesen aportar información acerca del mecanismo de localización de esta proteína en los *speckles* nucleares. Los resultados de los experimentos de *pull down* identificaron la asociación de la proteína nucleolar NOLC1 con los dominios FF4/FF5 de TCERG1; esta asociación no tenía lugar con los dominios mutados (Fig. I11) (Sánchez-Hernández 2013).

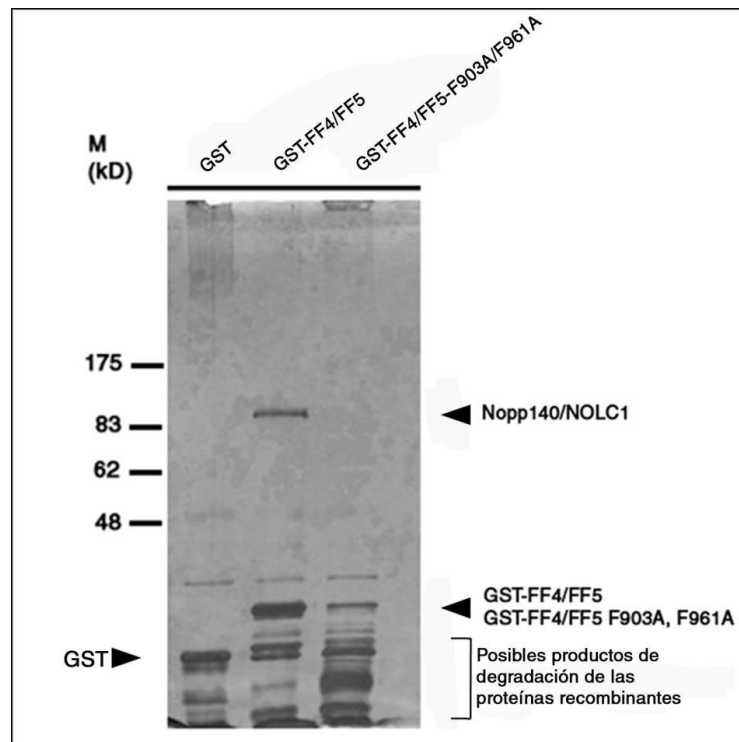


Figura I11. Ensayos de pull down usando los dominios FF4/FF5 WT y sin mutar.. Los precipitados se analizaron por SDS-PAGE y tinción de plata. El análisis de estos geles reveló la presencia de una proteína de aproximadamente 83 kDa asociada de forma específica a los dominios FF4/FF5 WT (GST-FF4/FF5) y no a los dominios mutados (GST-FF4/FF5-F903A/F961A). La banda se identificó por espectrometría de masas como Nopp140/NOLC1. Imagen obtenida de la tesis doctoral de Noemí Sánchez Hernández (Sánchez-Hernández 2013).

Objetivos

Los objetivos de esta tesis doctoral son:

1. Estudio de la asociación entre TCERG1 y NOLC1. Identificación de la región de interacción en TCERG1.
2. Estudio de la asociación entre TCERG1 y los CBs.
 - a. Estudio de la localización de TCERG1 en los CBs.
 - b. Estudio del efecto de la inhibición de TCERG1 en los CBs y otros cuerpos nucleares.
 - c. Estudio del efecto de una expresión saturante de TCERG1 en los CBs y otros cuerpos nucleares.
 - d. Estudio del efecto de NOLC1 y coilina en la localización de TCERG1, en los CBs y en otros cuerpos nucleares.
3. Estudio del papel de TCERG1 en la biogénesis de los snRNPs.
 - a. Estudio de la formación de los snRNPs al alterar la expresión de TCERG1.
 - b. Estudio de la asociación de TCERG1 con factores implicados en la biogénesis de los snRNPs.

Material y métodos

1. Plásmidos

Los plásmidos de expresión eucariota que se mencionan a continuación han sido descritos en trabajos anteriores del grupo: pEFBOST7-TCERG1 (Suñé and Garcia-Blanco 1999); pEFBOS-EGFP y pEFBOS/EGFP/T7-TCERG1[1-1098] (Sánchez-Álvarez et al. 2010); los mutantes de pEFBOS/EGFP/T7-TCERG1 y pEFBOST7-TCERG1 (Sánchez-Hernández et al. 2012; Sánchez-Hernández 2013); pcDNA3-EGFP (Addgene) y pcDNA3/EGFP/TCERG1(1–1098) (Muñoz-Cobo et al. 2017); pEFBOS/GFP/T7-SF1 (Becerra et al. 2015). El vector de expresión eucariota pEFBOST7-NOLC1 fue obtenido por métodos de PCR convencional utilizando la pareja de oligonucleótidos Nolc1T7FwXbaI y Nolc1T7RvBstBI con las dianas XbaI y BstBI. Los vectores de expresión bacteriana pGEX2TK-FF4/FF5 and pGEX2TK-FF4/FF5-F903A/F961A han sido también descritos previamente (Sánchez-Hernández 2013).

Para la generación de las líneas HEK293T establemente silenciadas mediante la tecnología CRISPR/Cas9, se utilizaron los plásmidos KN21792G1 para la transfección de la gRNA1, KN21792G2 para la transfección de la gRNA2, GE100003 para la transfección de la guía control con una secuencia al azar que no reconoce ningún gen conocido y el plásmido KN21792D como donador de resistencia a la puromicina para la selección de los clones (Origene).

2. Cultivos celulares y transfecciones

2.1. Cultivos celulares

Todas las líneas celulares que se detallan a continuación se cultivaron en condiciones estándares idénticas: 37°C en atmósfera de CO₂ del 5% y humedad del 95%.

Las células HEK293T y las células HeLa se cultivaron en medio de cultivo DMEM (Dulbecco's modified Eagle Medium, Gibco) con una concentración de 4,5 g/L de glucosa, suplementado con suero bovino fetal (Gibco) al 10%, penicilina/estreptomicina (100 U y 0.1 mg/ml, respectivamente) (Gibco) y 4 mM L-glutamina (Gibco). Las líneas celulares de HEK293T establemente silenciadas obtenidas mediante la tecnología CRISPR/Cas9 (1AC4 y 2AC2), así como la línea control (*Scramble*), se cultivaron con el medio descrito anteriormente adicionando 1,5 µg/ml de puromicina.

La línea celular de neuroblastoma SH-SY5Y se cultivó en medio DMEM con una concentración de 4,5 g/L de glucosa, suero fetal bovino al 10% (v/v) y penicilina/estreptomicina (100 U y 0,1 mg/ml, respectivamente). Para su cultivo se utilizaron placas revestidas con colágeno (BD biosciences). Para ello, se diluyó el colágeno a una concentración de 100 µg/ml en ácido acético 0,02 N. La solución se añadió a las placas que se incubaron entre 2 y 18 h a 37 °C en atmósfera húmeda con un 5% de CO₂. Posteriormente, se lavaron las placas dos veces con agua miliQ autoclavada, y se dejaron secar. Las células SH-SY5Y establemente silenciadas mediante shRNAs fueron descritas anteriormente (Muñoz-Cobo et al. 2017) y se mantuvieron con el mismo medio de cultivo adicionando 1 µg/ml de puromicina (Invitrogen) como antibiótico de selección.

Las líneas celulares HAP1 (Horizon Discovery) se mantuvieron en medio de cultivo IMDM (*Iscove's Modified Dulbecco's Medium*, Gibco) suplementado con un 10% de suero bovino fetal y penicilina/estreptomicina (100 U y 0,1 mg/ml, respectivamente).

La línea celular U2OS GFP-SmB fue cedida por el Dr. Angus I. Lamond (*Centre for Gene Regulation and Expression - School of Life Sciences, University of Dundee, Dundee, Scotland*) y mantenidas en medio de cultivo DMEM con una concentración de 4,5 g/L de glucosa, suplementado con suero bovino fetal (Gibco)

al 10%, penicilina/estreptomicina (100 U y 0.1 mg/ml, respectivamente) y con 150 mg/ml de higromicina B (Invitrogen) y 15 mg/ml de blasticidina-HCl (Gibco) como antibióticos de selección. Para la inducción de la expresión de GFP-SmB en estas células se añadieron 10 ng/ml de doxiciplina (Sigma). Las células se recogieron a las 18 h.

2.2. Transfecciones

En los experimentos de inmunoprecipitación usando construcciones etiquetadas con T7 y GFP se sembraron células HEK293T en placas de 100 mm de diámetro (BD Falcon) 24 h antes de la transfección. Cuando las células alcanzaron aproximadamente un 70-80% de confluencia fueron transfectadas con las cantidades de ADN correspondientes a cada plásmido mediante el método de precipitación del fosfato cálcico [$\text{Ca}_3(\text{PO}_4)_2$]. Brevemente, 500 μl de la solución A (250 mM CaCl_2) fueron mezclados con del ADN y posteriormente mezclados con 500 μl de la solución B (50 mM HEPES, 1,5 mM NaHPO_4 , 140 mM NaCl; pH ajustado a 7,05) durante 2 min para ser luego añadidos al cultivo. Después de aproximadamente 40 h de la transfección las células fueron recogidas y procesadas.

En los estudios de inmunofluorescencia y experimentos de sobre-expresión para análisis por *western blot* y RT-PCR se sembraron las células HEK293T o HeLa en placas de 6 pocillos de 35 mm de diámetro (BD Falcon), directamente o sobre cubreobjetos adecuados para microscopía confocal, aproximadamente 20 h antes de la transfección. Cuando la confluencia de las células fue del 50-60% para los experimentos de inmunofluorescencia y del 60-70% para los experimentos de sobre-expresión, se transfectaron con las correspondientes cantidades de plásmidos utilizando $\text{Ca}_3(\text{PO}_4)_2$. Las células transfectadas sobre los cubreobjetos se recogieron a las 20 h posteriores a la transfección. Las células para el análisis por

western blot y RT-PCR se suplementaron con medio fresco 24 h después de la transfección y se recogieron aproximadamente a las 40 h posteriores a la transfección.

En el caso de los experimentos con siRNAs en células HEK293T se sembraron aproximadamente 9×10^4 células por pocillo de 35 mm, directamente o sobre cubreobjetos adecuados para microscopía confocal, aproximadamente 16 h antes de la transfección. Cuando las células alcanzaron aproximadamente un 30-40% de confluencia se llevó a cabo la transfección de los siRNAs a una concentración final de 60 nM utilizando el reactivo Lipofectamine 2000 (Invitrogen) siguiendo las indicaciones del fabricante. Las células se recogieron a las 48 h posteriores a la transfección. En el caso de los experimentos usando células U2OS, los siRNA (35 nM) se transfectaron utilizando el reactivo lipofectamine RNAiMAX (Invitrogen) siguiendo las indicaciones del fabricante.

2.3. Generación de líneas establemente silenciadas mediante CRISPR/Cas9

Las células HEK293T fueron transfectadas con 1 μ g de cada plásmido que contiene la guía o la secuencia *Scramble* junto con 1 μ g del plásmido donador. Transcurridas 48 h se comprobó que la transfección había funcionado mediante análisis por *western blot* utilizando el anticuerpo anti-DDK, que reconoce la enzima Cas9 expresada (Fig. M1). Las células se mantuvieron durante 23 días hasta que se procedió a la selección mediante la adición de puromicina (1 μ g/mL) durante 5 días. Para la obtención de clones, las células se sembraron mediante dilución máxima (0,5 célula/pocillo) en placas de 96 pocillos y se seleccionaron a través del análisis de la expresión de la proteína silenciada por *western blot*.

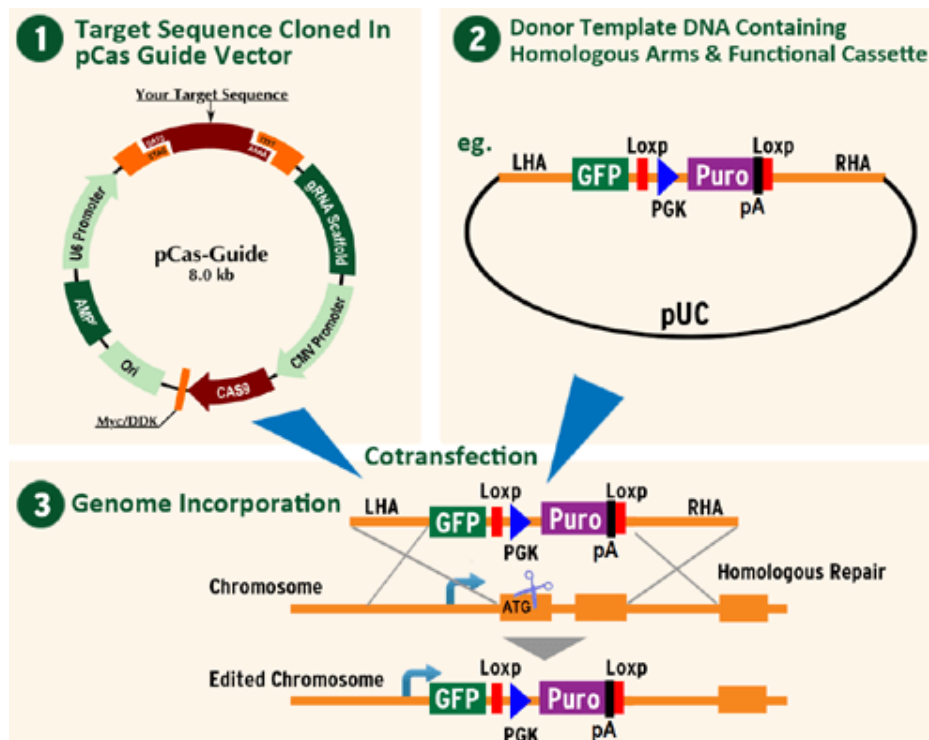


Figura M1. Esquema del kit de Origene utilizado para la generación de líneas estables mediante el sistema CRISPR/Cas9. En el esquema se representan los plásmidos guías (*pCas-Guide*), que contienen las secuencias guías dirigidas contra el gen TCERG1, y el plásmido donador (pUC) que contiene la secuencia que codifica el gen de resistencia a la puromicina

3. Procedimientos de biología celular

3.1. Inmunofluorescencia y microscopía confocal

Para los estudios de inmunofluorescencia, las células HEK293T, HeLa, HAP1 y U2OS cultivadas sobre cubreobjetos en placas de 6 pocillos fueron fijadas con paraformaldehído (PFA) (Sigma) al 3,5% en PBS 1X, 20mM HEPES (pH 7,4) durante 45 min a 4°C. Tras tres lavados con PBS 1X, las células se permeabilizaron con una solución de Triton-X100/PBS 1X (Sigma) al 0,5% durante 5 min a temperatura ambiente y se lavaron nuevamente tres veces con PBS 1X. A

continuación, fueron bloqueadas en una solución de albúmina de suero bovino (BSA, Sigma) al 2,5% en PBS 1X a 4°C durante al menos 16 h. Tras dos lavados con PBS 1X conteniendo 0,1% de BSA y un lavado con PBS 1X a temperatura ambiente, se procedió al tratamiento con anticuerpos. Se utilizó una misma solución de incubación y lavado para los anticuerpos primarios y secundarios (0,1% de BSA en PBS 1X), con la dilución correspondiente de cada anticuerpo. Los anticuerpos primarios se incubaron durante 1 h a temperatura ambiente en condiciones de humedad, y tras cinco lavados se procedió a incubar el anticuerpo secundario en las mismas condiciones. Por último se realizaron cinco lavados con 0,1% BSA/PBS 1X y tres lavados con PBS 1X, para posteriormente montar las preparaciones utilizando como medio de montaje Prolong Gold Antifade DAPI (Molecular Probes).

En el caso de los estudios de inmunofluorescencia usando células SH-SY5Y cultivadas en cubreobjetos tratados con colágeno, se fijaron primero con paraformaldehído al 4% y sacarosa (Sigma) al 4% en PBS (pH 7,4) durante 20 min a 37°C. A continuación, estas células se permeabilizaron con metanol frío al 90% durante 5 min y se bloquearon con 10% BSA /0,5% Triton X-100 en PBS 1X durante 1 h a temperatura ambiente. Los anticuerpos primarios se incubaron a la dilución adecuada en BSA/PBS 1X al 10% durante 2 h a temperatura ambiente o durante 16 h a 4°C en condiciones de humedad. Tras tres lavados con Tx-100/PBS 1X al 0,5% durante 10 min cada uno, las células se incubaron con los anticuerpos secundarios diluidos en la misma solución de Tx-100/PBS 1X al 0,5% durante 2 h a temperatura ambiente en condiciones de humedad. Por último, se lavaron tres veces con BSA/PBS 1X al 0,1% y tres veces con PBS 1X durante 10 min por lavado y se prepararon para microscopía utilizando como medio de montaje Prolong Gold Antifade (Molecular Probes).

Las imágenes fueron adquiridas con un microscopio confocal invertido Leica SP5. Las imágenes se tomaron simultáneamente en los casos en que se realizaba

un doble marcaje con EGFP o el fluorocromo Alexa 647, excitados con el láser de helio-neón (HeNe) de 633 nm, y el fluorocromo Alexa 488, excitado con el láser de argón a 488 nm al 20% de potencia. Para el doble marcaje con el fluorocromo Alexa 550, excitado con el láser HeNe 543, y EGFP o el fluorocromo 647, las imágenes se tomaron de forma secuencial. En ambos casos el espectro de detección para cada canal se ajustó para evitar el paso de señal de un canal a otro. Las imágenes se tomaron en condiciones de no saturación con un objetivo 63x, resolución 1024 x 1024 y un diámetro de ranura confocal (*pinhole*) de 1 μ m. Las imágenes se analizaron con el software Las AF (Leica). Posteriormente las imágenes fueron procesadas con el software Adobe Photoshop CS3 v10.0.

3.2. Sincronización del ciclo celular

Las células HEK293T fueron cultivadas sobre cubreobjetos en placas de 6 pocillos que se incubaron con hidroxIUrea (ThermoFisher) a una concentración de 0,5 mM durante 18 h para obtener una población sincronizada en fase G1. La población de células sincronizadas en fase S se obtuvo incubando con hidroxIUrea a una concentración de 0,5 mM durante 18 h y posteriormente se cambió el medio por DMEM fresco durante 3 h antes de recoger las células. Los cubreobjetos se recogieron para su análisis mediante microscopía confocal. Para los análisis de citometría de flujo, el resto de células del pocillo se recogieron mediante tripsinización en tubos de poliestireno y se lavaron dos veces con PBS 1X centrifugando a 500 G a 4°C. Las células fueron fijadas con etanol al 70% a 4°C durante al menos 15 min. Tras un par de lavados con PBS 1X, las células fueron incubadas durante 20 min a 37°C en oscuridad, en una solución de PBS 1X que contenía 100 μ g/ml de RNasa A (Roche) y 40 μ g/ml de yoduro de propidio (PI) (Sigma). Las muestras fueron analizadas por citometría de flujo usando un

FACScalibur (BD). Los datos fueron procesados utilizando el software FlowJo v7.6.5.

4. Manipulación y análisis de proteínas

1.1. Purificación de IgGs totales del suero anti-TCERG1

Fracciones de inmunoglobulinas tipo G (IgGs) totales fueron obtenidas de sueros inmuno-reactivos contra TCERG1 mediante cromatografía de afinidad en una columna con agarosa conjugada a proteína A (GE Healthcare Life Science). Se preparó una columna de cromatografía "Poly-Prep" (Bio-Rad) de 1 ml de resina de agarosa lavada previamente en 0,1M Tris-HCl pH 8, y equilibrada con 10 volúmenes de este mismo tampón. A continuación se cargaron dos veces a una velocidad inferior a 0,3 ml/min, 1 ml de suero anti-TCERG1 diluido un décimo en Tris 1 M pH 8 para igualar el pH y la concentración salina de la muestra con la columna. La columna se lavó secuencialmente a una velocidad mayor de 0,5 ml/min con 10 ml de 0,1M Tris-HCl pH 8 y 10 ml de 0,01 M Tris-HCl pH 8, respectivamente. Posteriormente, las IgGs se eluyeron de la columna añadiendo 3 ml de 0,1 M Glicina pH 3 a una velocidad menor de 0,3 ml/min y se recogieron en fracciones de aproximadamente 270 μ L en tubos donde previamente se habían añadido 30 μ L de 1M Tris-HCl pH 8 para neutralizar el pH ácido del eluido. Después de medir la concentración de proteínas de las fracciones a partir de la absorbancia a 280 nm, las fracciones con una concentración de proteínas mayor a 0,3 mg/mL se juntaron y se sometieron a diálisis con dos etapas secuenciales de 2 y 16 h respectivamente usando 2 litros de PBS con azida sódica al 0,1% en cada una.

4.1. Preparación de extractos de proteínas

Las células fueron recogidas de la placa de cultivo y centrifugadas durante 5 min a 800 G a 4°C. El precipitado obtenido se lisó con tampón de lisis T7 frío (20mM HEPES, pH 7,9, 150 mM NaCl, 5 mM EDTA, 1% Nonidet fosfato-40 (NP-40), 1mM ditioneitol (DTT) y 1mM fenilmetil-sulfonil-fluoruro (PMSF) e inhibidores de proteasas (*Complete*, Roche) o tampón de lisis RIPA frío (50 mM TrisHCl pH 7,5, 1% NP-40, 0,05% dodecil sulfato sódico (SDS), 1 mM EDTA, 150 mM NaCl, 0,5 % Sodium deoxicolate, 1mM DTT, 1mM PMSF e inhibidores de proteasas (*Complete*) durante 30 min a 4°C. Posteriormente, el lisado se centrifugó a 1600 G durante 5 min a 4°C y el sobrenadante utilizado como extracto total.

4.1. Análisis por *western blot*

Las muestras a analizar se diluyeron en tampón de carga Laemli 2x (0.2% azul de bromofenol, 10% glicerol, 0.1 M TrisHCl pH 6.8, 2% dodecil sulfato sódico (SDS), 50 mM ditioneitol (DTT)) y se calentaron 5 min a 99°C. Después de cargar las muestras en geles de poliacrilamida (SDS-PAGE) al 10%, la electroforesis se llevó a cabo en tampón Laemli (25mM Tris, 192mM de glicina y 0,1% SDS) a un voltaje continuo de 150 V. Terminada la electroforesis, las proteínas se transfirieron a membranas de nitrocelulosa (Amersham Biosciencies) mediante un sistema húmedo de transferencia en una solución que contenía 25 mM Tris, 192 mM de glicina y 20% de metanol durante 1 h a 4°C con una intensidad constante de 200-300 mA. Las membranas se trataron con una solución de bloqueo (5% leche desnatada, 0,1% Tween-20, PBS 1X) durante 1h como mínimo. Las incubaciones con los distintos anticuerpos se realizaron en esta misma solución de bloqueo, ajustándose las diluciones y condiciones de incubación para cada anticuerpo primario. Los anticuerpos secundarios conjugados con peroxidasa (PerkinElmer Life Science) fueron utilizados a una dilución 1:5000 durante 1h a temperatura

ambiente. Los lavados de los distintos anticuerpos se realizaron con la solución 0,1% Tween-20, PBS 1X. La reacción de revelado se efectuó con un reactivo comercial (*Western Lightning Chemiluminiscence Reagent Plus*, Perkin Elmer).

En los casos cuyo objetivo era cuantificar los niveles de expresión de proteínas, se utilizaron anticuerpos secundarios acoplados a fluoróforos en el rango infrarrojo cercano (IR) y su detección se realizó usando el sistema *Odyssey CLx*. El análisis de cuantificación se realizó con el programa *Image Studio*.

Para los casos en los que se requirió la reutilización de las membranas, se procedió a realizar un lavado extensivo en PBS-Tween (0.1% Tween-20 en PBS), 2 incubaciones de 15 min a 50°C en solución ST (62.5mM TrisHCl pH 6.8, 2% (p/v) SDS, 100 mM β -mercaptoetanol (Sigma)), y 6 lavados de 10 min en PBS-Tween. Las membranas así tratadas se bloquearon de nuevo y se incubaron con nuevos anticuerpos de la forma en que se ha descrito previamente.

4.2. Precipitación mediada por resina acoplada a GST (Ensayos de *pull down*)

Para los ensayos de *pull down*, 6 μ g de cada proteína de fusión fueron pre-aclarados en rotación en 40 μ l de resina de agarosa acoplada a glutatión en 800 μ l de tampón de unión de GST (20 mM Tris HCl, pH 7,9, 180 mM KCl, 0,2 mM EDTA, 0,05% NP-40, 0,5 mM PMSF, 1 mM DTT) que contenía BSA a 1 mg/mL. A la vez, 200 μ l de extracto nuclear (NE) de células HeLa (CilBiotech) fueron diluidos en 800 μ l de buffer de unión GST y pre-aclarados en 40 μ l de resina. Tras 2 h, la proteína unida a la resina fue incubada con los extractos nucleares pre-aclarados durante 16 h en rotación a 4°C. A continuación y tras 5 lavados de 5 min cada uno con tampón de lavado GST (20 mM Tris HCl, pH 7,9, 180 mM KCl, 0,2 mM EDTA, 0,1% NP-40, 0,5 mM PMSF, 1 mM DTT), se eluyeron las proteínas y los

complejos unidos añadiendo 30 μ l de tampón de carga SDS 2X (0,2% de azul de bromofenol, 100 mM DTT, 4% SDS, 20% glicerol y 0,1M Tris HCl, pH 6,8) durante 10 min a 99°C. Los eluidos fueron resueltos en geles de SDS-PAGE al 10% para ser posteriormente analizados mediante tinción de plata.

4.3. Tinción de plata

Los geles a analizar se sumergieron en agitación suave durante 30 min en solución de fijado (40% metanol, 10% ácido acético). A continuación, el gel fue lavado durante 5 min en la solución A (10% etanol, 5% ácido acético) y sensibilizado mediante una solución de tiosulfato sódico ($\text{Na}_2\text{S}_2\text{O}_3$) al 0,01% durante 5 min. Tras varios pasos consecutivos de lavado con la solución A (durante 2 min), etanol al 10% (durante 5 min) y agua destilada (durante 5 min) se procedió a la tinción del gel con nitrato de plata (AgNO_3) al 0,1% durante 5 min en agitación y oscuridad. Por último, el revelado del gel se realizó en dos pasos. Primero, y durante 5 min, el gel se incubó en una solución de carbonato sódico (Na_2CO_3) al 10%. En segundo lugar, el gel se sumergió en una solución de Na_2CO_3 al 2% y formaldehído al 0,01% durante 20 s. Transcurrido este tiempo se realizó un breve lavado con agua destilada y se volvió a añadir la misma solución hasta la aparición de bandas. La tinción se detuvo realizando lavados sucesivos con agua destilada.

4.4. Ensayos de inmunoprecipitación

4.4.1. Inmunoprecipitación de proteínas endógenas

Para los ensayos de inmunoprecipitación se sembraron 2×10^6 células en placas de 100 mm de diámetro (Falcon). Después de aproximadamente 40 h las

células fueron lisadas en 500 μ l de tampón de lisis RIPA o tampón T7 frío. Tras eliminar los restos celulares por centrifugación se mezcló un décimo de cada muestra con tampón Laemli 2x y se guardó convenientemente para su análisis por *western blot* como muestra del extracto celular total (*Input*). O bien, se usaron 50 μ l de extractos nucleares (NE) de células HeLa (CilBiotech) para por ensayo que se diluyeron en tampón de lisis RIPA o T7 frío. 40 μ l de una mezcla de resina de Proteína A + G (Millipore) se incubaron con las muestras durante 1 h en rotación a 4°C. La resina se eliminó mediante centrifugación y las muestras se incubaron con los anticuerpos correspondientes y con IgGs de ratón, de conejo o de cobaya, como controles negativos de la inmunoprecipitación, durante toda la noche en rotación a 4°C. Al día siguiente, las muestras se incubaron con 50 μ l de una mezcla de resina de Proteína A + G durante 2 h a 4°C en rotación. Después de 5 lavados con el tampón de lisis utilizado anteriormente, las muestras se eluyeron añadiendo 50 μ l de tampón Laemli 2X e incubando durante 5 min a 99 °C en agitación (1050 rpm) para su posterior análisis mediante *western blot*.

4.4.2. Inmunoprecipitación de proteínas expresadas transitoriamente con el epítipo T7.

Para cada ensayo se sembraron 2×10^6 de células HEK293T en placas de 100 mm de diámetro (BD Falcon). Al día siguiente y cuando la confluencia alcanzaba el 90%, las células se transfectaron con la cantidad correspondiente de cada plásmido. Aproximadamente 40 h después de la transfección las células fueron recogidas en 2 ml de PBS 1X frío (4°C) y lisadas en 500 μ l de tampón de lisis RIPA o T7 frío. Tras eliminar los restos celulares por centrifugación se mezcló un décimo de cada muestra con tampón Laemli 2x y se guardó convenientemente para su análisis por *western blot* como muestra del extracto celular total (*whole cellular extract*, WCE). Para el tratamiento con RNAsa, los lisados celulares se trataron con

0,1 mg/ml de RNAsa A (Roche) durante 1 h a 4°C en rotación antes de la incubación con la agarosa. Los lisados fueron diluidos hasta 1 mL con el buffer de lisis correspondiente, e incubados durante 4-6 h a 4°C en rotación con 50 µl de agarosa covalentemente unida a un anticuerpo monoclonal anti-T7 (Novagen). Tras 5 lavados con el tampón correspondiente, la resina se incubó durante 5 min a 99°C en agitación y se centrifugó durante 1 min a máxima velocidad a temperatura ambiente. El sobrenadante se analizó mediante SDS-PAGE al 10% y *western blot*.

4.4.3. Inmunoprecipitación de proteínas expresadas transitoriamente con epítomos diferentes a T7.

Se sembraron 2×10^6 células HEK293T en placas de 100 mm de diámetro. Cuando la confluencia alcanzó el 90%, se transfectó la cantidad correspondiente de cada plásmido y a las 40 h post-transfección, las células se recogieron en 2 mL de PBS 1X frío y se lisaron en 500 µl de tampón de lisis T7 o RIPA frío. Se guardó un décimo de cada muestra (WCE) con tampón Laemli 2X para su posterior análisis mediante *western blot*. Los lisados fueron diluidos hasta 1 mL en el correspondiente tampón de lisis e incubados durante 1 h a 4°C en rotación con 40 µl de una mezcla de resina de Proteína A + G, previamente lavada. A continuación, se desechó esta resina y el sobrenadante se incubó con los anticuerpos correspondientes durante toda la noche a 4°C en rotación. Al día siguiente, la muestra se incubó con 50 µl de una mezcla de resina de agarosa de proteína A + G, previamente lavada durante 2 h a 4°C en rotación. Tras 5 lavados con el tampón correspondiente, la resina de agarosa se incubó durante 5 min a 99°C en agitación y se centrifugó durante 1 min a máxima velocidad a temperatura ambiente. El sobrenadante se analizó mediante SDS-PAGE al 10% y *western blot*.

4.4.4. Inmunoprecipitación de RNA (RIP)

Células HEK293T y HAP1 fueron sembradas en placas de 100 mm de diámetro y recogidas al alcanzar el 100% de confluencia con su propio medio de cultivo. Se fijaron con formaldehído (Sigma) al 1% durante 10 min a temperatura ambiente en agitación y posteriormente se glicina al 0,125 M durante 5 min a temperatura ambiente en agitación para detener el entrecruzamiento. Las células se centrifugaron, se lavaron con PBS 1X frío y se lisaron con 1 ml de tampón RIPA durante 30 min a 4°C. A continuación se sonicaron 3 veces durante 20 s en hielo (vibra-cell, Sonics materials) y se centrifugó a velocidad máxima durante 10 min a 4°C. Una fracción del sobrenadante se conservó para la extracción de RNA y la extracción de proteína como *input*. El resto del sobrenadante se incubó con 40 µl de una mezcla de resina de Proteína A + G conteniendo 100 µg/ml de tRNA durante 2 h a 4°C en rotación. La resina se eliminó por centrifugación y el sobrenadante se incubó con anticuerpos específicos y anti-IgGs a 4°C en rotación toda la noche. A continuación, la muestra se incubó con 50 µl de la resina de agarosa de la mezcla de Proteína A + G durante 2 h a 4°C en rotación. La resina fue lavada seis veces con tampón RIPA durante 10 min a 4°C (por lavado) en rotación y posteriormente se resuspendieron en 100 µl de tampón 50 mM TrisHCl pH 7.0, 5 mM EDTA, 10 mM DTT, 1% SDS, y se incubaron a 70°C durante 45 min. La muestra se dividió en dos partes, una para la extracción del RNA con peqGOLD Trifast según las indicaciones del fabricante y la otra se mezcló con Laemli 5X para su posterior análisis mediante *western blot*.

5. Procedimientos de biología molecular

5.1. Extracción de RNA y análisis por RT-qPCR

El RNA total se extrajo con peqGOLD Trifast (peQlab) según las indicaciones del fabricante y se digirió con ADNasa (Invitrogen) siguiendo las indicaciones de la casa comercial. A continuación, 1 µg de RNA se retro-transcribió utilizando qScript cDNA Supermix (Quanta Bioscience) siguiendo las indicaciones del fabricante, que contenía una mezcla de oligo (dT), *random primers* y la retro-transcriptasa *qScript*. El producto de la reacción resultante se diluyó cinco veces para su análisis. La cuantificación por PCR a tiempo real (qPCR), se llevó a cabo utilizando *iQ SYBR Green Supermix* (Bio-Rad) en la estación *iCycler termal cyler station* (Bio-Rad) con oligonucleótidos específicos. Como control interno de cada experimento se midieron los niveles de gliceraldehído-3-fosfato deshidrogenasa (GAPDH) utilizando los oligonucleótidos GAPDH-Fw/GAPDH-Rv. Los experimentos se realizaron tres veces como mínimo. Los valores están representados como:

$$(E \text{ "gen problema"})^{\Delta CT \text{ ("gen problema")}} / (E \text{ GAPDH})^{\Delta CT \text{ (GAPDH)}}$$

Donde E = eficiencia de la PCR y $\Delta CT = CT \text{ muestra control} - CT \text{ muestra problema}$.

6. Análisis estadístico

Para el análisis estadístico utilizado en esta tesis, se usaron los programas Prism 5.0 software (GraphPad) y Microsoft Office Excel. Se aplicó el test t-Student para comparar las medias entre muestras. Los valores P estadísticamente

significativos se representan en las gráficas mediante asteriscos (*, $P < 0,05$; **, $P < 0,01$; ***, $P < 0,005$).

7. Anticuerpos y oligonucleótidos

7.1. Anticuerpos

| Anticuerpo | Especie | Origen | Dilución |
|----------------------|---------|--|-------------------------|
| α TCERG1 | Conejo | Suero (Suñé and Garcia-Blanco 1999) | IF 1:1000 WB 1:60000 |
| α TCERG1 | Cobaya | IgG purificado | WB 1:10000 |
| α Coilin (F7) | Ratón | Santa Cruz Biotechnology (sc-55594) | WB 1:1000 IF 1:250 |
| α NOLC1 | Conejo | Abcam (ab106324) | WB 1:1000 IF 1:500 |
| α CDK9 | Conejo | Santa Cruz Biotechnology (sc-484) | WB 1:2000 |
| α T7 | Conejo | Bethyl Laboratories (#A190-116A) | WB 1:20000 IF 1:1000 |
| α Sm (Y12) | Conejo | Thermo (#MS-450) | WB 1:1000 IF 1:200 |
| α SMN | Ratón | Santa Cruz Biotechnology (sc-32313) | WB 1:500 IF 1:200 |

| Anticuerpo | Especie | Origen | Dilución |
|---|---------|---|-----------|
| αGFP | Ratón | Roche (11 814 460 001) | WB 1:5000 |
| αGAPDH | Ratón | Santa Cruz Biotechnology (sc-365062) | WB 1:2000 |
| αSRSF2 | Ratón | Sigma (S4045) | IF 1:4000 |
| αLsm11 | Conejo | Cedido por el Dr. Joseph G. Gall (CRNAegie Institution for Science, Baltimore, MD) | IF 1:1000 |
| αDDK | Ratón | Origene (TA50011-100) | WB 1:2000 |
| αTMG | Ratón | Merk Millipore (MABE302) | IF 1:500 |
| αNucleolin | Conejo | Abcam (ab22758) | IF 1:2000 |
| αCRM1 | Conejo | Santa Cruz Biotechnology (sc-5596) | WB 1:1000 |
| αCBP80 | Conejo | Santa Cruz Biotechnology (sc-48803) | WB 1:1000 |
| αInt11 | Conejo | Cedido por el Dr. Edouard Bertrand (IGMM-CNRS, Montpellier, Francia) | WB 1:1000 |
| αIgGRabbit- HRP | Cabra | Perkin Elmer Life Science (#NEF812001EA) | WB 1:5000 |
| αIgGMouse- HRP | Cabra | Perkin Elmer Life Science (#NEF822001EA) | WB 1:5000 |
| αIgGGuineaPig- HRP | Cabra | Jackson ImmunoResearch | WB 1:5000 |

| Anticuerpo | Especie | Origen | Dilución |
|---|---------|----------------------------------|-----------|
| | | 106-035-003 | |
| αIgGRabbit-Alexa488 | Cabra | Molecular Probes (#A-11008) | IF 1:500 |
| αIgGMouse-Alexa647 | Cabra | Molecular Probes (#A-21235) | IF 1:500 |
| αIgGRabbit-Alexa550 | Cabra | Bethyl Laboratories (A120-201D3) | IF 1:500 |
| αIgGMouse-800CW IRDye | Cabra | LI-COR (925-32210) | WB 1:5000 |
| αIgGRabbit-800CW IRDye | Cabra | LI-COR (925-32211) | WB 1:5000 |

WB: *Western blot*

IF: Inmunofluorescencia

7.2. Oligonucleótidos

Clonaje NOLC1

| Nombre oligonucleótido | Secuencia 5' - 3' |
|------------------------|------------------------------|
| Nolc1T7FwXbaI | GGGTCTAGAATGGCGGACGCC |
| Nolc1T7RvBstBI | GGGTTCGAATCACTCGCTGTCAAACCTT |

qPCR

| Nombre oligonucleótido | Secuencia 5' - 3' |
|------------------------|--------------------------------|
| GAPDH Fw | GATGGGGAAGGTGAAGGTCG |
| GAPDH Rv | GGGTCATTGATGGCAACAATATC |
| U1snRNA Fw | GATACCATGATCACGAAGGTGGTT |
| U1snRNA Rv | CACAAATTATGCAGTCGAGTTTCC |
| U2snRNA Fw | TTTGGCTAAGATCAAGTGTAGTATCTGTTC |
| U2snRNA Rv | AATCCATTTAATATATTGTCCTCGGATAGA |
| U4snRNA Fw | GCGCGATTATTGCTAATTGAAA |
| U4snRNA Rv | AAAAATTGCCAATGCCGACTA |
| U5snRNA Fw | GGTTTCTCTTCAGATCGCATAAATC |
| U5snRNA Rv | CTCAAAAAATTGGGTTAAGACTCAGA |
| U6snRNA Fw | GCTTCGGCAGCACATATACTAAAAT |
| U6snRNA Rv | ACGAATTTGCGTGTATCCTT |
| U1-pre Fw | ACTGCGTTCGCGCTTTCCC |
| U1-pre Rv | GCAGGCGACATGTTACTTCC |
| U2-pre Fw | AACATAGGTACACGTGTGCCACGG |
| U2-pre Rv | ACAAATAGCCAACGCATGCGGGGC |

| Nombre oligonucleótido | Secuencia 5' - 3' |
|------------------------|--------------------------|
| Coilin Fw | CTTGAGAGAACCTGGGAAATTTG |
| Coilin Rv | GTCTTGGGTCAATCAACTCTTTCC |

siRNA

| Nombre | Secuencia 5' - 3' |
|-----------|-------------------------|
| siGFP | CUACAACAGCCACAACGUC |
| siTCERG1 | GGAGUUGCACAAGAUAGUU |
| siControl | CAGUCGCGUUUGCGACUGG |
| siNOLC1 | AAAUUGAGGUGGAUUCACGAGUU |
| siCoilin | GAGAGAACCUGGGAAUUUU |

gRNA CRISPR

| Nombre | Secuencia 5' - 3' |
|---------------------|----------------------|
| gRNA 1 (KN21792G1) | GAGTGAACGATTCAACCCGG |
| gRNA 2 (KN217192G2) | TCTCGTGAGCAGAGATGTGA |

Resultados

1. TCERG1 interacciona con la fosfoproteína nucleolar NOLC1

1.1. Identificación de la secuencia necesaria para la interacción entre TCERG1 y NOLC1

Como se ha descrito en la Introducción, la proteína TCERG1 se localiza en la periferia de los *speckles* nucleares, y se caracteriza por tener 6 repeticiones en tándem de dominios FF. Estudios anteriores de nuestro laboratorio identificaron que en los dominios FF4 y FF5 se encontraba la región necesaria para que se diera lugar esta localización. En este mismo estudio, se demostró que los residuos altamente conservados de fenilalaninas 903 y 961 eran requeridos para la localización de la proteína TCERG1 a la periferia de los *speckles* nucleares.

Basándonos en estos datos, y con el fin de identificar proteínas de unión a los dominios FF4/FF5 que mediaran la localización de TCERG1 a los *speckles*, se generaron construcciones que contenían los dominios FF4/FF5 de TCERG1 y estos dominios con la doble mutación F903A/F961A, fusionados a la proteína GST (GST-FF4/FF5 y GST-FF4/FF5-F903A/F961A, respectivamente) para realizar ensayos de *pull down*. En estos ensayos, se identificó una banda asociada a los dominios FF4/FF5 que estaba ausente en las muestras que usaban como cebo la doble mutación F903A/F961A. Dicha banda se aisló de geles y se analizó mediante espectrometría de masas (MALDI TOF-TOF) en la Unidad de Proteómica del Centro Nacional de Biotecnología (CSIC). El análisis de las masas obtenidas identificó a la fosfoproteína NOLC1 (Nopp140) en las muestras procedentes del *pull-down* con GST-FF4/FF5. La identidad de la proteína NOLC1 se comprobó en las muestras originales usadas en el *pull-down* mediante *western blot* (WB) y anticuerpos específicos (Sánchez-Hernández 2013). Para confirmar estos

resultados e iniciando esta tesis doctoral, se repitieron de nuevo los ensayos de *pull-down* y las muestras resultantes se analizaron en la Unidad de Proteómica del IPBLN-CSIC. El análisis de las masas obtenidas, en tres experimentos independientes, identificó de nuevo a la fosfoproteína NOLC1 (Nopp140) en las muestras procedentes del *pull-down* con GST-FF4/FF5.

Para confirmar la interacción entre TCERG1 y NOLC1 se realizaron experimentos de inmunoprecipitación en células utilizando distintas construcciones de TCERG1 acopladas al epítipo T7 (Figura R1-A). La transfección de la construcción pEFBOST7-TCERG1 (WT) en células HEK293T y el análisis de los inmunoprecipitados, obtenidos mediante el uso de anticuerpos anti-T7, demostró la presencia de NOLC1 en los mismos (Figura R1-B, TCERG1). Con el fin de comprobar que la interacción con NOLC1 era mediada por los dominios FF4/FF5, se utilizó una construcción que contenía el extremo amino terminal de TCERG1 fusionado a los dominios FF4 y FF5 (Figura R1-A, Nt-FF4/FF5), y otra construcción que contenía la secuencia de TCERG1 completa excepto los dominios FF4 y FF5 (Figura R1-A, TCERG1- Δ FF4/ Δ FF5). El análisis de los inmunoprecipitados obtenidos al transfectar transitoriamente dichas construcciones reveló que NOLC1 se asociaba a TCERG1 en presencia de los dominios FF4/FF5 (Figura R1-B, ver los carriles señalados como Nt-FF4/FF5 y TCERG1- Δ FF4/ Δ FF5). Como control positivo, utilizamos el anticuerpo específico para la quinasa CDK9, cuya interacción con TCERG1 ha sido descrita previamente (Sánchez-Álvarez et al. 2006).

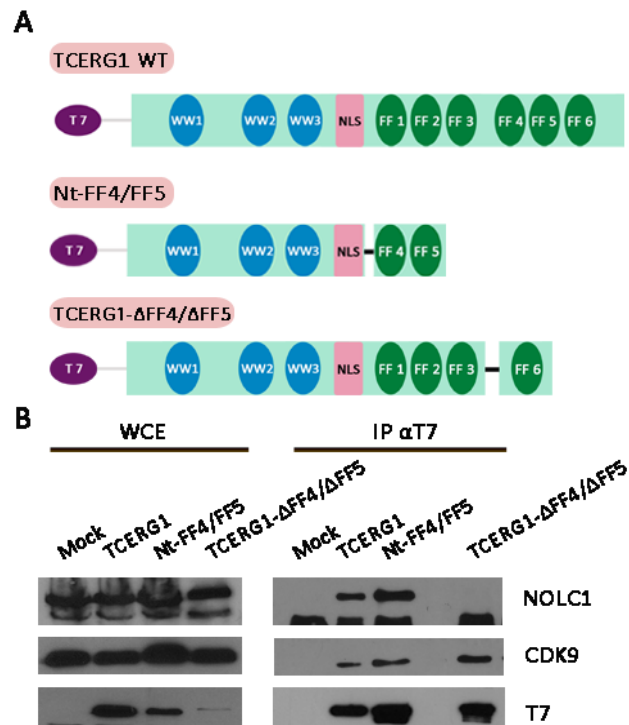


Fig R1. Los dominios FF4 y FF5 de TCERG1 son necesarios para la asociación entre TCERG1 y NOLC1. A. Representación esquemática de las construcciones de TCERG1 utilizadas en el ensayo de inmunoprecipitación. Se indican los dominios WW, los dominios FF y la región que contiene la NLS fusionados al epítipo T7. B. Células HEK293T fueron transfectadas transitoriamente con los plásmidos que expresan las distintas construcciones de TCERG1 unidas al epítipo T7 o un vector vacío como control negativo (Mock). Fracciones del extracto celular total (WCE) fueron analizadas directamente por *western blot* o inmunoprecipitadas con anticuerpos específicos contra T7, seguida de una separación por SDS-PAGE y análisis por WB utilizando anticuerpos específicos para NOLC1, CDK9 y T7.

Para comprobar la importancia de los residuos de fenilalanina 903 y 961 de los dominios FF4/FF5 en la interacción entre TCERG1 y NOLC1, se llevaron a cabo otra serie de inmunoprecipitaciones. Para ello, se utilizaron construcciones con el extremo amino terminal de TCERG1 unido a los dominios FF4 y FF5 incluyendo o no las mutaciones en los residuos de fenilalanina (Figura R2-A, Nt-FF4/FF5 y Nt-FF4/FF5 [F903, 961A]). Usando resina de agarosa acoplada al anticuerpo anti-T7 procedimos a inmunoprecipitar las construcciones sobre-expresadas en células

HEK293T y a analizar los inmunoprecipitados por WB. Se observó que la proteína NOLC1 interaccionaba con la construcción de TCERG1 que incluía los dominios FF4 y FF5 sin las fenilalaninas mutadas (Figura R2-B, Nt-FF4/FF5). Por lo que se deduce que las fenilalaninas 903 y 961 de los dominios FF4/FF5 de TCERG1 son necesarias para que se dé lugar la interacción con NOLC1.

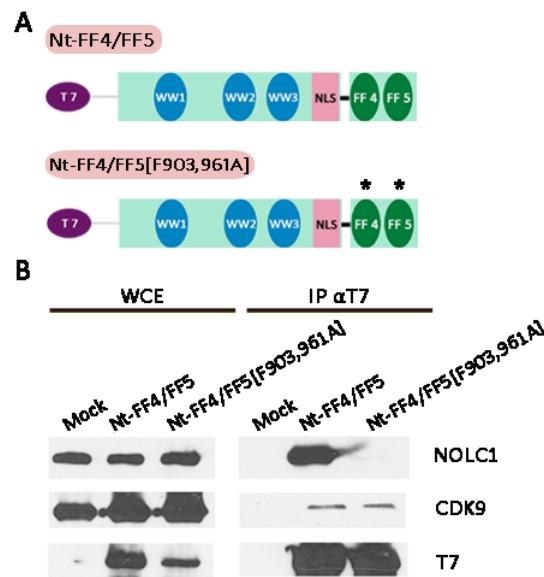


Figura R2. Las fenilalaninas 903 y 961 de los dominios FF4 y FF5 son necesarias para la asociación entre TCERG1 y NOLC1. A. Representación esquemática de las construcciones de TCERG1 utilizadas en el ensayo de inmunoprecipitación. Se indican los dominios WW, los dominios FF y la región que contiene la NLS central de la proteína fusionados al epítipo T7. Los asteriscos representan las mutaciones puntuales en esos dominios. B. Los plásmidos indicados se transfectaron transitoriamente en células HEK293T. Un vector vacío unido a T7 se utilizó como control negativo (Mock). Fracciones de extracto celular total (WCE, *whole cell extract*) fueron analizadas directamente por WB o inmunoprecipitadas con anticuerpos específicos contra T7, seguida de una separación por SDS-PAGE y análisis por *western blot* utilizando anticuerpos específicos para NOLC1, CDK9 y T7.

Para confirmar estos resultados, repetimos el ensayo de inmunoprecipitación, pero en este caso, transfectando de forma transitoria el

vector de expresión pEFBOST7-NOLC1 donde se encuentra el DNA que codifica para la proteína NOLC1 fusionado al epítipo T7. Los extractos inmunoprecipitados, utilizando anticuerpo anti-T7 unido covalentemente a resina de agarosa, fueron analizados mediante WB usando anticuerpos específicos dirigidos contra TCERG1. Se pudo comprobar que TCERG1 se encontraba presente en los inmunoprecipitados de las muestras procedentes de la sobre-expresión de NOLC1 (Figura R3, IP T7-NOLC1). Por tanto, confirmamos las observaciones previas sobre la interacción entre NOLC1 y TCERG1.

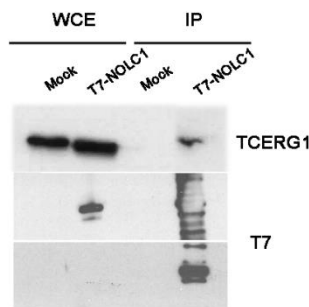


Figura R3. NOLC1 sobre-expresado interacciona con TCERG1. El plásmido T7-NOLC1 o un vector vacío unido a T7 se transfectaron transitoriamente en células HEK293T. Fracciones de extracto celular total (WCE, *whole cell extract*) fueron analizadas directamente por *western blot* o inmunoprecipitadas con anticuerpos específicos contra T7, seguida de un análisis por *western blot* utilizando anticuerpos específicos para TCERG1 y T7. En los paneles inferiores se muestran dos exposiciones diferentes del análisis utilizando el anticuerpo anti-T7.

A continuación, quisimos confirmar estos resultados con las proteínas TCERG1 y NOLC1 endógenas. Para ello, utilizamos un anticuerpo policlonal anti-TCERG1 y su control, suero preinmune de cobaya, para inmunoprecipitar extractos nucleares de células HeLa. El análisis posterior mediante *western blot* se realizó utilizando anticuerpos específicos contra NOLC1. Los resultados obtenidos revelaron la presencia de NOLC1 en los inmunoprecipitados procedentes de las muestras tratadas con anticuerpos anti-TCERG1 (Figura R4, IP TCERG1),

indicando una asociación entre las proteínas endógenas que confirmaba las observaciones previas sobre la interacción entre estas dos proteínas. Con estos resultados podemos concluir que TCERG1 se asocia a la fosfoproteína NOLC1.

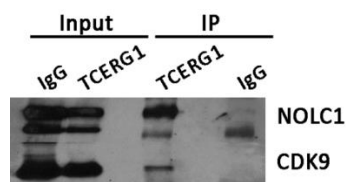


Figura R4. Las proteínas TCERG1 y NOLC1 endógenas interactúan entre sí. Extractos nucleares de células HeLa se inmunoprecipitaron con un anticuerpo específico contra TCERG1 o bien utilizando suero preinmune de cobaya como control negativo (IgG). Una fracción de la muestra (Input) se guardó para realizar el análisis mediante *western blot* junto con los inmunoprecipitados utilizando anticuerpos específicos contra NOLC1 y CDK9, este último utilizado como control positivo.

1.2. La interacción entre TCERG1 y NOLC1 está mediada por RNA.

Con el fin de seguir estudiando la interacción observada entre TCERG1 y la fosfoproteína NOLC1 y dada la función de TCERG1 en el procesamiento del pre-mRNA, nos pareció interesante comprobar si esta asociación estaba mediada a través de RNA. Para ello, células HEK293T transfectadas con la construcción T7-TCERG1 se recogieron y se lisaron con buffer de lisis RIPA para posteriormente, tratar los extractos proteicos con 0,1 mg/mL de RNAsa A durante 1 h a 4°C. Como control positivo, células transfectadas con la misma construcción se sometieron a los mismos tratamientos de temperatura pero sin la adición de la RNAsa y, a continuación, se realizaron ensayos de inmunoprecipitación utilizando anticuerpo

anti-T7 covalentemente unido a resina de agarosa A. El análisis mediante *western blot* indicó que la asociación entre TCERG1 y NOLC1 es dependiente de RNA, ya que no se observó interacción entre estas dos proteínas en las muestras tratadas con RNAsa (Figura R5, +/- RNasa).

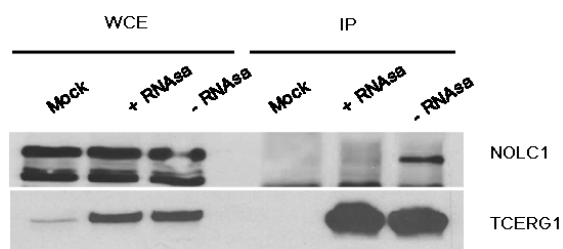


Figura R5. La asociación entre TCERG1 y NOLC1 es dependiente de RNA. Células HEK293T se transfectaron con T7-TCERG1 y un plásmido T7 vacío como control negativo (Mock). Los extractos proteicos se trataron con RNAsa A, posteriormente inmunoprecipitados con anti-T7 y analizados mediante *western blot* utilizando los anticuerpos NOLC1 y TCERG1.

2. TCERG1 se asocia a los cuerpos de Cajal

2.1. NOLC1 interacciona con coilina.

NOLC1 se describió originalmente como una chaperona que se desplaza del nucléolo al citoplasma (Thomas Meier and Blobel 1992). La localización principal de NOLC1 es en el componente fibrilar denso del nucléolo y además se ha observado que se encuentra en las regiones nucleares denominadas cuerpos de Cajal (*Cajal Bodies, Coiled Bodies, CBs*) (Isaac, Yang, and Thomas Meier 1998). Nuestros análisis mediante inmunofluorescencia indirecta, corroboró los estudios anteriores sobre la localización de esta proteína en el nucleolo y en los CBs, ya que al utilizar un anticuerpo específico contra coilina, considerada el marcador de los

cuerpos de Cajal, NOLC1 muestra una acumulación evidente en estos orgánulos nucleares (Fig. R6-A). Estos resultados sugieren una asociación espacial entre coilina y NOLC1.

Estudios anteriores observaron que existía una interacción *in vitro* entre las proteínas NOLC1 y coilina en el sistema del doble híbrido en levaduras (Isaac, Yang, and Thomas Meier 1998). Para comprobar si esta interacción tenía lugar en líneas celulares las muestras procedentes de la inmunoprecipitación de extractos nucleares transfectados transitoriamente con el plásmido pEFBOST7-NOLC1 del experimento mostrado en la Figura R3 se analizaron mediante WB utilizando un anticuerpo monoclonal específico para coilina. Los resultados obtenidos mostraron que NOLC1 se asocia con coilina en células (Figura R6-B), corroborando los datos anteriormente mencionados.

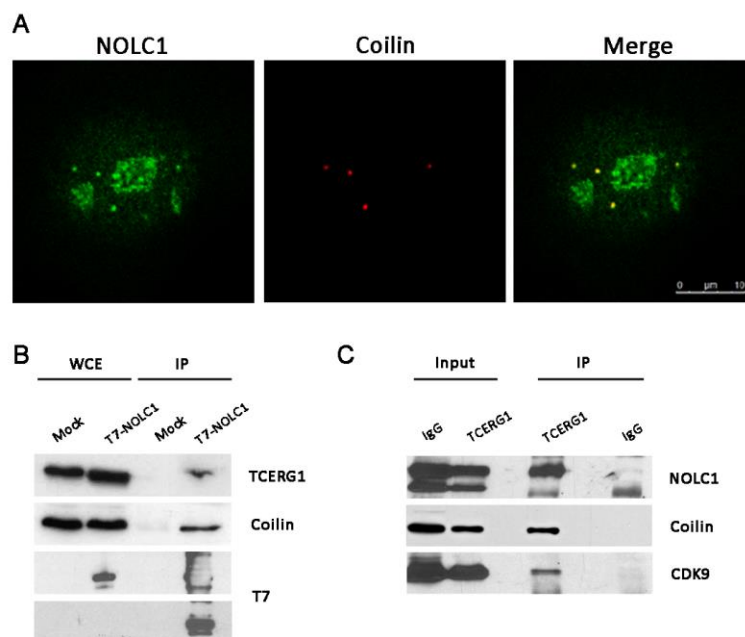


Figura R6. Asociación entre TCERG1, NOLC1 y coilina. A. El uso de la inmunofluorescencia indirecta utilizando anticuerpos que reconocen las proteínas endógenas NOLC1 (verde) y coilina (rojo), así como la superposición de ambas señales (Merge), revela que existe una asociación entre

ellas en los CBs. **B.** Las muestras inmunoprecipitadas utilizando un anticuerpo frente al epítipo T7 procedentes de extractos transitoriamente transfectados con T7-TCERG1 y control, se analizaron mediante *western blot* con los anticuerpos indicados. **C.** TCERG1 se asocia a coilina. Análisis de las muestras procedentes del experimento mostrado en la Figura R4 por WB usando anticuerpos específicos para NOLC1, coilina y CDK9.

2.2. TCERG1 se asocia a coilina

A la vista de los resultados anteriores y dado que TCERG1 interacciona con NOLC1, nos preguntamos si TCERG1 y coilina se asocian. Para ello se utilizaron las muestras procedentes de la inmunoprecipitación de proteínas endógenas usando anticuerpo anti-TCERG1 (Figura R4) para realizar un ensayo de WB con anticuerpos específicos para coilina. Los resultados mostraron que las proteínas endógenas TCERG1 y coilina interaccionan entre sí (Figura R6-C) y que, por tanto, existe una asociación entre ellas.

Estos resultados nos llevaron a seguir estudiando la asociación entre estas dos proteínas. Para ello, llevamos a cabo ensayos de inmunofluorescencia utilizando anticuerpos específicos dirigidos contra TCERG1 y coilina en células HEK293T. Las imágenes obtenidas de estos experimentos mostraron, en primer lugar, un patrón de localización similar al esperado según los estudios anteriores. TCERG1 se encuentra difusamente localizado en el nucleoplasma celular acumulándose en los *speckles* nucleares (Sánchez-Álvarez et al. 2006) y coilina se localiza específicamente en los CBs (Figura R7-A, paneles TCERG1 y coilin). La superposición de imágenes (Figura R7-A, panel Merge) muestra mínima colocalización entre ambas proteínas. Un análisis en mayor profundidad utilizando el programa Fiji/ImageJ (Figura R7-B) que marca zonas de colocalización de los elementos de imagen (píxeles) en color blanco, y el análisis semicuantitativo de la relación espacial entre las distribuciones de TCERG1 y

coilina (Figura R7–C), sugieren que TCERG1 no se acumula en los CBs, sino que se encuentra localizado en la periferia de estos orgánulos nucleares.

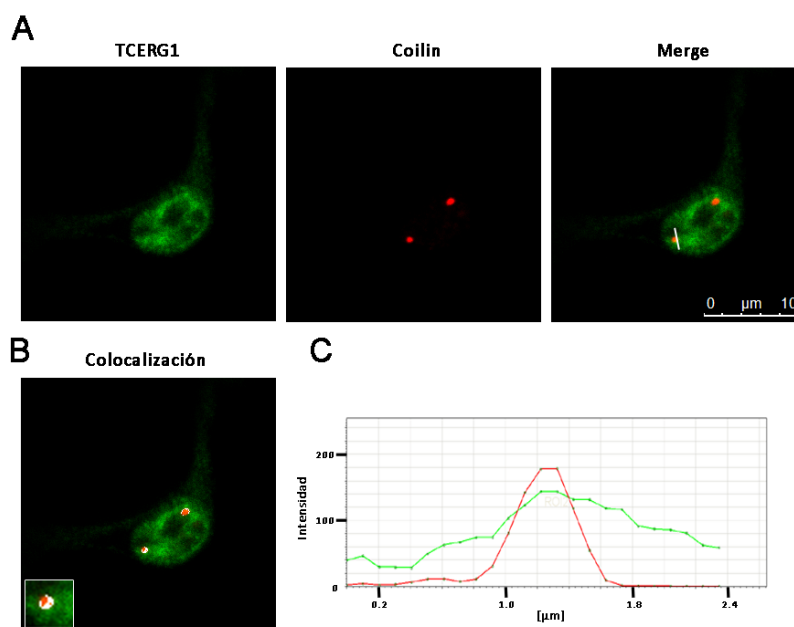


Figura R7. Análisis de la colocalización de TCERG1 y coilina A. Se muestran imágenes de células HEK293T marcadas doblemente con anticuerpos que detectan TCERG1 (verde) y coilina (rojo), así como la superposición de ambas (Merge). La línea dibujada en la imagen superpuesta indica la zona de cuantificación para el análisis en C. B. Análisis de colocalización mediante el programa Fiji/ImageJ. Los píxeles blancos indican las zonas de colocalización. C. Análisis semicuantitativo de la intensidad de las señales cuya región aparece indicada en las imágenes superpuestas que se muestran en A.

A continuación decidimos analizar el patrón de localización de TCERG1 y coilina a lo largo del ciclo celular. Para ello, utilizamos el compuesto hidroxurea para obtener poblaciones sincronizadas en fase G1 en células HEK293T (ver Material y Métodos) (Fig. R8-B, panel derecho). El estudio de la colocalización (fig. R8-A) y el análisis semicuantitativo de la intensidad de las señales (Fig. R8-B, panel izquierdo), indicó un incremento de la colocalización entre TCERG1 y coilina en las células que se encontraban en la fase G1. Al retirar la hidroxurea a

las células durante 3 h, se obtuvo una población de células sincronizadas en fase S (Fig R8-D panel derecho). Realizando un análisis por microscopía confocal de la colocalización, se observó un menor solapamiento de las señales de TCERG1 y coilina durante esta fase (Fig. R8-C). El análisis semi-cuantitativo corroboró este resultado (Fig. R8-D, izquierda). De estos resultados concluimos que hay un incremento de la colocalización entre TCERG1 y coilina durante la fase G1, que es la etapa donde la tasa de transcripción de los snRNAs es más alta y, por ello, se encuentran en mayor densidad en la periferia de los CBs (Wang et al. 2016).

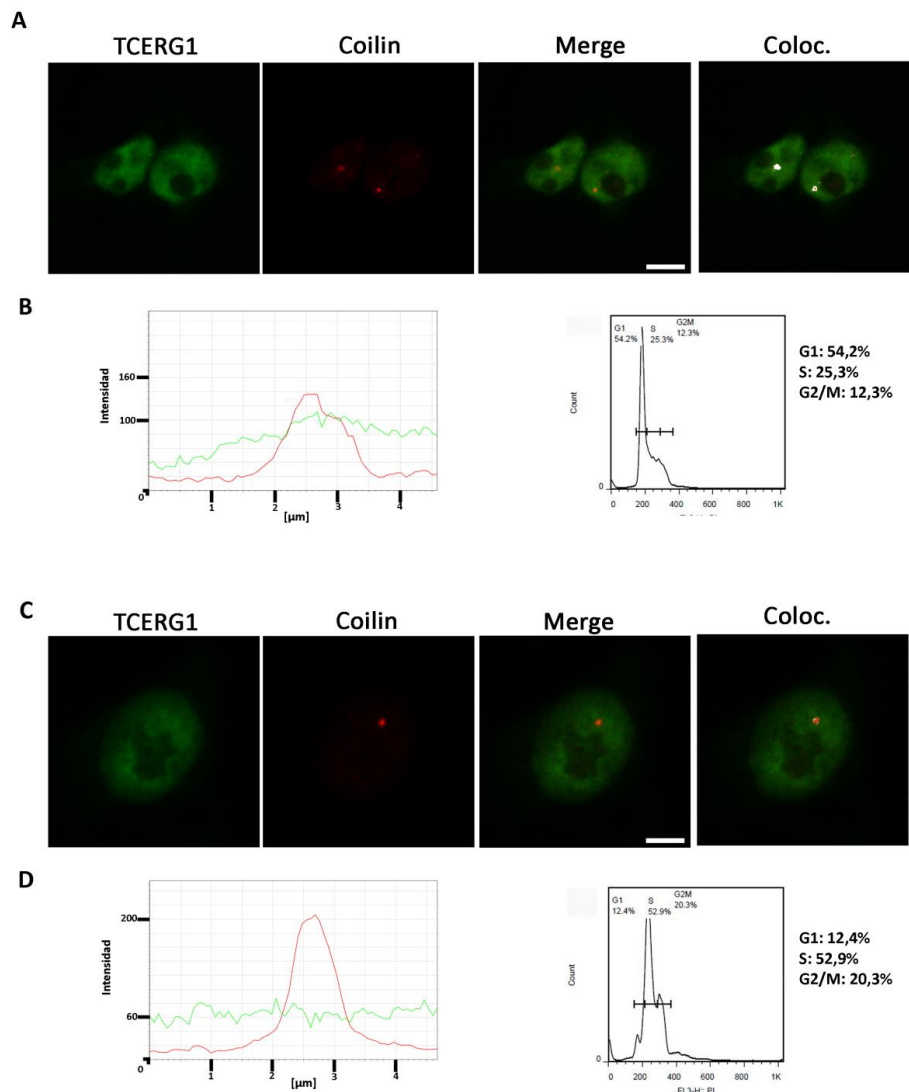


Figura R8. Localización de TCERG1 y coilina durante el ciclo celular. **A y C.** Imágenes de IF donde se muestra la tinción de TCERG1 (verde) y coilina (rojo), la superposición de ambas señales (Merge) y el análisis de colocalización mediante el programa Fiji/ImageJ (Coloc.). Los píxeles blancos indican las zonas de colocalización. **B.** Análisis semi-cuantitativo de la intensidad de las señales (panel izquierdo). En el panel derecho se muestra el histograma de citometría de flujo de células HEK293T sincronizadas en fase G1 marcadas con yoduro de propidio (IP). Se muestra el porcentaje de células que se encuentran en cada fase del ciclo celular. **D.** En el panel izquierdo se muestra el análisis semi-cuantitativo de la intensidad de las señales. En el panel derecho se muestra el histograma de citometría de flujo de células HEK293T sincronizadas en fase S marcadas con yoduro de propidio (IP). Se muestra el porcentaje de células que se encuentran en cada fase del ciclo celular.

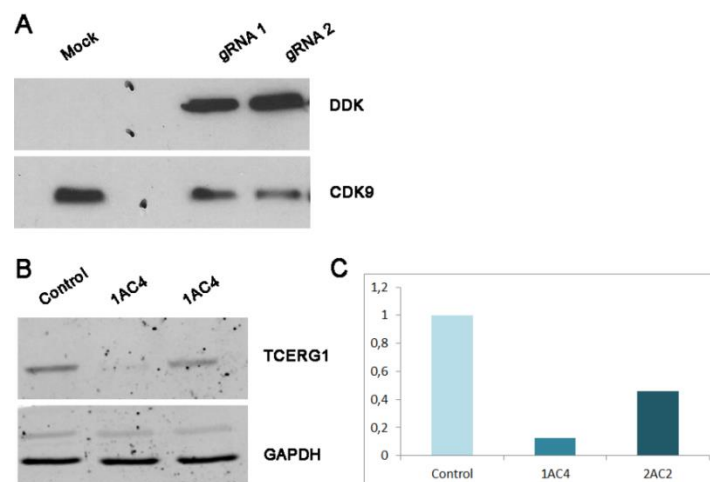
2.3. Efecto de la inhibición de TCERG1 en la localización de coilina.

2.3.1. Generación de líneas HEK293T silenciadas para TCERG1 mediante CRISPR/Cas9

Con el objetivo de estudiar la relación entre TCERG1 y los CBs, generamos líneas celulares silenciadas establemente para TCERG1 mediante la tecnología CRISPR/Cas9. Para ello, utilizamos dos guías diferentes dirigidas a dos regiones distintas del gen *TCERG1*; la guía 1 (gRNA1) reconoce el primer exón y la guía 2 (gRNA2) reconoce el primer intrón de dicho gen. Cada una de las guías junto con el plásmido donador, que contiene la resistencia a Puromicina, fueron transfectadas en células HEK293T. La eficiencia de la transfección se comprobó utilizando un anticuerpo anti-DDK (Figura R9-A). A los 23 días posteriores a la transfección, las células fueron seleccionadas mediante la adición de puromicina. Asimismo y para la generación de la línea control, se utilizó una guía (*Scramble*) que contenía una secuencia que no reconocía ningún gen conocido humano y se transfectó junto con el plásmido donador, procediendo de forma análoga a las

anteriores. Se generaron diferentes clones y mediante análisis por *western blot* se seleccionó un clon por cada guía utilizada, obteniendo el clon 1AC4 procedente de la gRNA1 y el clon 2AC2 procedente de la gRNA2 (Figura R8-B). La cuantificación de la expresión de la proteína TCERG1 indicó una expresión de TCERG1 menor del 20% en la línea 1AC4 y de aproximadamente el 50% en la línea 2AC2 (Figura R9-C). A pesar de nuestros esfuerzos, no fue posible obtener líneas *knockout* para TCERG1.

A continuación se realizó un marcaje por inmunofluorescencia utilizando un anticuerpo anti-TCERG1 junto con un anticuerpo específico de la forma fosforilada del factor esencial de *splicing* SRSF2 (anteriormente conocido como SC35), el cual se utiliza frecuentemente como marcador de los *speckles* nucleares. Se observó una disminución en la intensidad de la señal correspondiente a TCERG1 en las líneas celulares 1AC4 y 2AC2 respecto a la línea control, siendo menor en la línea 1AC4 corroborando los resultados obtenidos mediante *western blot*. El marcaje simultáneo de TCERG1 y SRSF2 mostró un solapamiento de las señales en los *speckles* nucleares (Figura R9-D) tal y como se esperaba de trabajos anteriores (Sánchez-Álvarez et al. 2006).



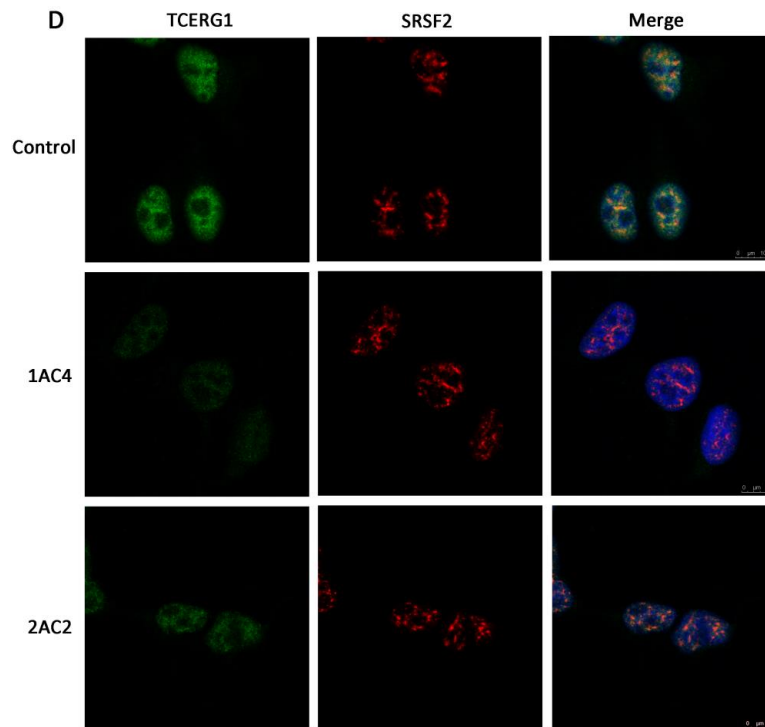


Figura R9. Análisis de las líneas silenciadas establemente para TCERG1 mediante CRISPR/Cas9. **A.** Análisis de la expresión de la enzima Cas9 en las células HEK293T transfectadas para la obtención de las líneas establemente silenciadas. Para ello, se utilizó el anticuerpo anti-DDK, que reconoce un epítipo de la enzima Cas9; la expresión de CDK9 se utilizó como control. Como control negativo se cargó una muestra de extractos de proteínas de células no transfectadas (Mock). **B.** Análisis de la expresión de la proteína TCERG1 en los clones seleccionados. La expresión de las proteínas en los extractos celulares de células control y silenciadas para TCERG1 se analizó por *western blot*, utilizando anticuerpos específicos; la expresión de GAPDH se utilizó como control. **C.** Cuantificación de la expresión de TCERG1 en los clones 1AC4 y 2AC2 respecto a las células control mediante *Image Studio*. **D.** Análisis de la localización de TCERG1 mediante ensayos de inmunofluorescencia en las líneas generadas. Se muestran imágenes marcadas doblemente con anticuerpos que detectan TCERG1 (verde) y SRSF2 (rojo), así como la superposición de ambas junto con la señal correspondiente al DAPI (azul) que marca los núcleos (Merge).

2.3.2. Efecto de la disminución de la expresión de TCERG1 sobre los CBs en células HEK293T.

Para estudiar el efecto de TCERG1 sobre los CBs, se procedió a realizar un ensayo de inmunofluorescencia de doble marcaje utilizando anticuerpos específicos para TCERG1 y coilina en las tres líneas celulares generadas (1AC4, 2AC2 y control). Este ensayo mostró que en situación de menor expresión de TCERG1 había un menor número de Cuerpos de Cajal y estos orgánulos eran de menor tamaño en las líneas silenciadas respecto a la línea control (Figura R9-A). Estas imágenes se adquirieron en las mismas condiciones de ganancia para ver las diferencias entre las líneas silenciadas y la línea control. Si nos fijamos en el marcaje de coilina, sólo se pueden observar los CBs en la línea control (Figura R-9, paneles etiquetados como coilin). Pero si modificamos las imágenes, amplificando la señal de rojo (Figura R-10, paneles etiquetados como coilin [sobre-expuesto]), podemos observar que los CBs en la línea control son más grandes y más intensos que los que se pueden ver en las líneas silenciadas al sobre-exponer la imagen. Con el fin de cuantificar estas observaciones, se obtuvieron y procesaron de manera independiente colecciones de imágenes para contar el número de CBs presentes en las tres líneas celulares utilizando el programa Fiji/ImageJ. En las líneas silenciadas de TCERG1, el número de CBs es menor que en las células control (Figura R10-B, panel izquierdo). Además, utilizando el mismo programa, se midió el tamaño de los CBs y se observó que los CBs encontrados en las líneas silenciadas eran de menor tamaño respecto a los de la línea control (Figura R10-B, panel derecho). Estos resultados sugieren que la disminución de la expresión de TCERG1 afecta la integridad de los CBs, medida por la presencia de coilina.

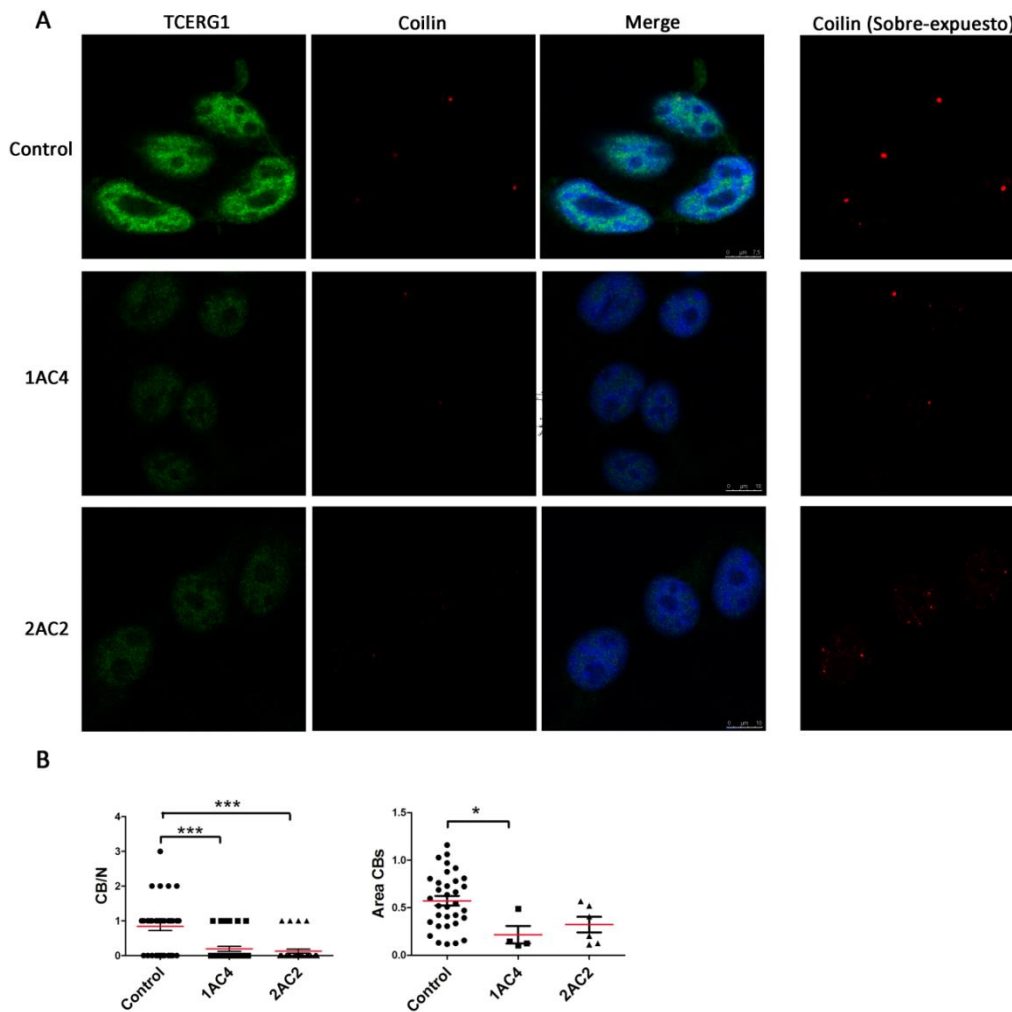


Figura R10. Efecto de la disminución de la expresión de TCERG1 en los CBs. **A.** Se muestran imágenes marcadas doblemente con anticuerpos que detectan TCERG1 (verde) y coilina (rojo), así como la superposición de ambas junto con la señal correspondiente al DAPI (azul) que marca los núcleos (Merge). En el panel derecho se muestran las imágenes del marcaje de cada línea celular con el anticuerpo frente a coilina, saturando la señal correspondiente al rojo (coilin [sobre-expuesto]). **B.** Cuantificación del número de CBs por núcleo (CB/N) en las tres líneas celulares y medida del tamaño de estos orgánulos (Area CBs). Poner número de núcleos medidos y estadística, como hacemos en el artículo.

2.3.3. Efecto de la disminución de la expresión de TCERG1 sobre los CBs en células HAP1.

A continuación quisimos confirmar los resultados anteriores usando otras líneas celulares. Las células de la línea humana HAP1 cuasi-haploide derivada de

la línea celular humana KBM-7. Se trata de células adherentes con morfología de fibroblasto, con un cariotipo haploide para todos los cromosomas excepto para 30 Mb del cromosoma 15 (Horizon). Mediante la tecnología CRISPR/Cas9, se obtuvo una línea (c012) donde la expresión de TCERG1 se encontraba muy disminuida (Figura R11-A). De nuevo, y como sucedió con las células HEK293T, no fue posible obtener líneas *knockout* para TCERG1, sugiriendo que esta proteína es esencial para la viabilidad celular. Los resultados de expresión mostraron en la línea c012 una banda ligeramente de mayor movilidad electroforética y de menor intensidad de señal respecto a la línea HAP1 control (Figura R11-A, c012). Estudios de inmunofluorescencia indirecta con anticuerpos anti-TCERG1 mostraron una menor presencia de proteína TCERG1 endógena en la línea c012, corroborando los resultados obtenidos mediante *western blot*. El marcaje simultáneo de SRSF2 y TCERG1 indicó una colocalización de ambas señales en gránulos nucleoplasmáticos (*speckles*) en la línea control, mientras el estudio de los perfiles lineales de intensidad indicó una localización más difusa de ambas proteínas en el nucleoplasma de la línea c012 (Figura R11-B).

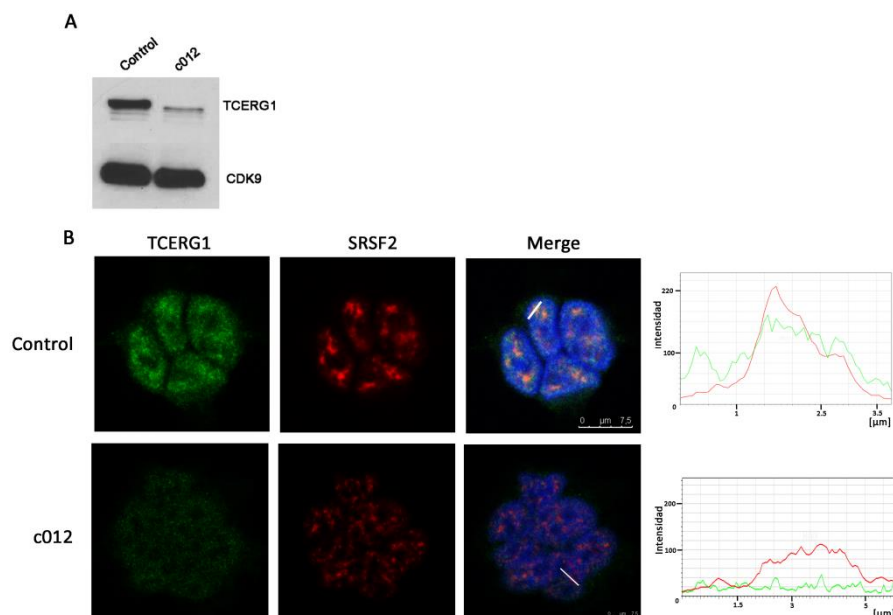


Figura R11. Estudio de la expresión y localización de TCERG1 en las líneas HAP1 control y c012. **A.** Análisis mediante *western blot* de la proteína TCERG1 en las líneas HAP1, tanto en la línea control como en la línea c012. El anticuerpo anti-CDK9 se usó como control. **B.** Se muestran imágenes de inmunofluorescencia de las células HAP1 doblemente marcadas con anticuerpos que detectan TCERG1 (verde) y SRSF2 (rojo), este último utilizado como marcador de *speckles* nucleares, así como la superposición de ambas señales (Merge). En los paneles derechos se representa el análisis semi-cuantitativo de la distribución de los perfiles lineales de las intensidades. Las líneas dibujadas en las imágenes superpuestas indican la zona de cuantificación en dicho análisis.

Para estudiar el efecto de la ausencia de TCERG1 en los CBs de las líneas HAP, se realizó un doble marcaje con anticuerpos que reconocían TCERG1 y coilina. De manera similar al estudio realizado en células HEK293T, en condiciones de menor expresión de la proteína TCERG1 se observó un menor número de cuerpos de Cajal (Figura R12-A, coilin). La cuantificación de los CBs indicó un menor número y tamaño de estos orgánulos en la línea c012 respecto a la línea control. Estos resultados corroboran los datos obtenidos con las células HEK293T y sugieren que la disminución de la expresión de TCERG1 afecta la integridad de los CBs.

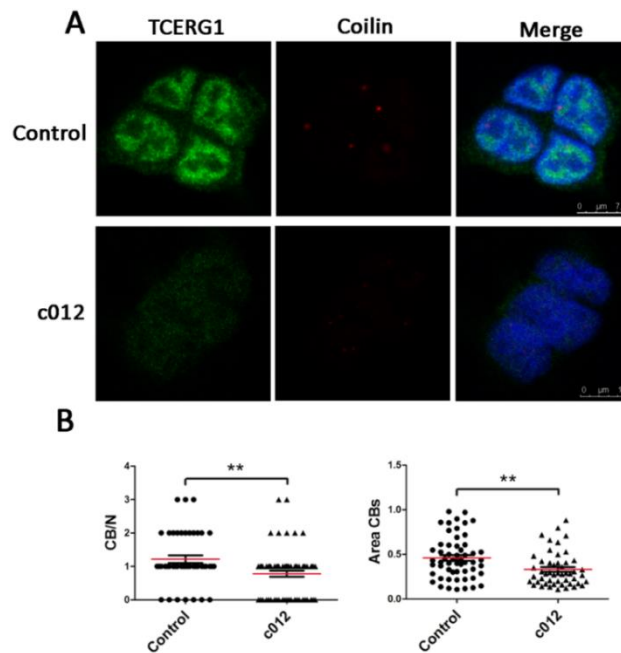


Fig R12. Efecto de la disminución de la expresión de TCERG1 sobre los CBs en las células HAP1. **A.** Se muestran imágenes de inmunofluorescencia indirecta con marcaje doble usando anticuerpos que detectan TCERG1 (verde) y coilina (rojo), así como la superposición de ambas, junto con la señal azul correspondiente al DAPI que marca los núcleos (Merge). **B.** Cuantificación del número de *Cajal Bodies* por núcleo (CB/N) y medida del tamaño de estos orgánulos (Area CBs) en las dos líneas celulares.

2.3.4. La ausencia de TCERG1 no afecta a la expresión de coilina.

A la vista de los resultados anteriores, se quiso comprobar si la expresión de coilina estaba siendo afectada por la disminución de la expresión de TCERG1. Para ello, se obtuvieron extractos de proteína de las líneas CRISPR HEK293T y HAP1 y se realizó un análisis mediante WB utilizando un anticuerpo específico para coilina. Los resultados indicaron que la expresión de coilina era similar en todas las líneas, tanto control como silenciadas (Figura R-13A-B). Estos resultados sugieren que la inhibición de TCERG1 afecta la localización de coilina en los CBs pero no a su expresión a nivel de proteína.

Para comprobar que la ausencia de TCERG1 no afectaba la expresión de coilina a nivel de RNA, se realizaron ensayos de RT-qPCR usando muestras de las líneas HEK293T. Se observó que el nivel de expresión de coilina era similar entre las células control y las líneas 1AC4 y 2AC2 (Fig R13-C). Estos resultados se corroboraron utilizando células HEK293T silenciadas transitoriamente mediante un RNA de interferencia que reconoce el mRNA del gen *TCERG1* (siTCERG1). Como en el caso anterior, la expresión del RNA mensajero de coilina fue similar entre la línea silenciada y la línea control (siGFP) (Fig R13-D).

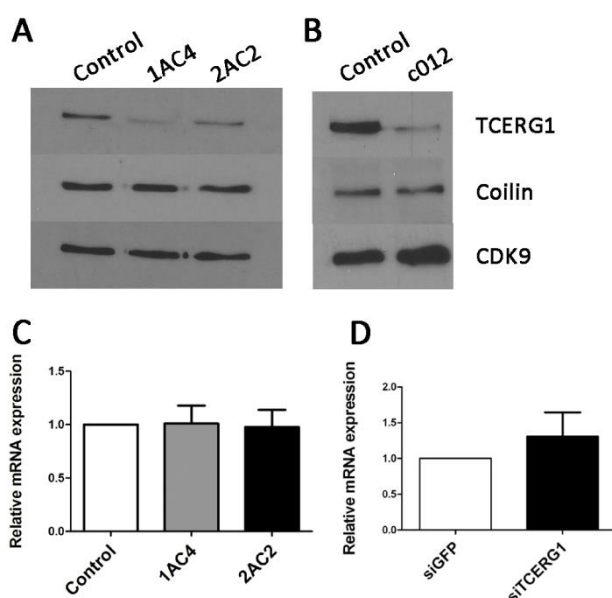


Figura R13. Análisis de la expresión de coilina en condiciones de una menor expresión de TCERG1. **A.** Análisis por WB de extractos de proteína de las líneas CRISPR/Cas9 de células HEK293T, tanto la línea control, como las líneas cuyo nivel de expresión de proteína de TCERG1 está disminuida (1AC4 y 2AC2). Las proteínas fueron detectadas utilizando anticuerpos específicos para TCERG1 y coilina. El anticuerpo anti-CDK9 se usó como control. **B.** El mismo análisis descrito en A se llevó a cabo con extractos de las líneas HAP1 control y c012. **C.** Representación gráfica de los datos obtenidos de la RT-qPCR del RNA obtenido de las líneas CRISPR/Cas9 de HEK293T. Las gráficas muestran los niveles de expresión relativa de coilina con respecto al control, considerado como 1. Los niveles de expresión fueron normalizados utilizando como control interno el gen

GAPDH. D. Representación gráfica de los datos obtenidos de la RT-qPCR de células HEK293T silenciadas con siTCERG1. Como control se utilizaron células transfectadas con siGFP. Las gráficas muestran los niveles de expresión relativa de coilina con respecto al control, considerado como 1. Los niveles de expresión fueron normalizados utilizando como control interno el gen GAPDH.

2.3.5. Efecto de la disminución de la expresión de TCERG1 sobre los CBs en células SH-SY5Y

En trabajos anteriores del laboratorio se había generado una línea celular de neuroblastoma (SH-SY5Y) establemente silenciadas de TCERG1 mediante un shRNA específico, denominada sh9, y una línea control usando un shRNA (pLKO.1) que no reconoce ningún gen de mamífero (Muñoz-Cobo et al. 2017). Estas líneas fueron utilizadas para comprobar si una menor expresión de TCERG1 afectaba a la localización de coilina en los CBs. Los resultados obtenidos mostraron unos resultados similares a los ya observados en las otras líneas celulares: un menor número de *Cajal bodies* en la línea SH-SY5Y establemente silenciada de TCERG1 (Fig. R14). Estos resultados corroboran que la expresión de TCERG1 afecta la integridad de los CBs.

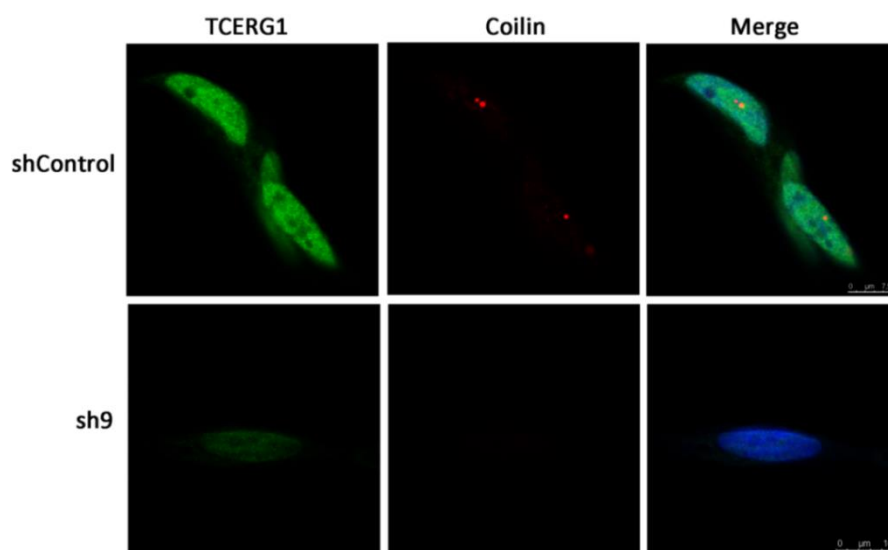


Figura R14. Efecto de la disminución de la expresión de TCERG1 sobre los CBs en las células SH-SY5Y. Se muestra un marcaje simultáneo de las proteínas endógenas TCERG1 (verde) y coilina (rojo) en células SH-SY5Y, tanto la línea control como la línea establemente silenciada de TCERG1 (sh9). Las imágenes de superposición de ambas señales (Merge) se muestran en los paneles derechos.

2.4. La disminución de la expresión de TCERG1 no afecta la localización de otros orgánulos nucleares.

Se quiso comprobar si otros orgánulos nucleares estaban siendo afectados por la disminución de la expresión de TCERG1.

2.4.1. *Speckles*

Como ya se ha mencionado, TCERG1 se encuentra en la periferia de los *speckles* nucleares. Por ello era especialmente relevante comprobar si una disminución de la expresión de TCERG1 podía afectar la integridad de estos orgánulos. Los experimentos realizados en células CRISPR HEK293T y HAP1 y ya expuestos en las figuras R9-D y R11-B, respectivamente, muestran cambios poco significativos en la integridad de los *speckles*, aunque se mostraron de menor tamaño y más difusamente distribuidos en las células c012.

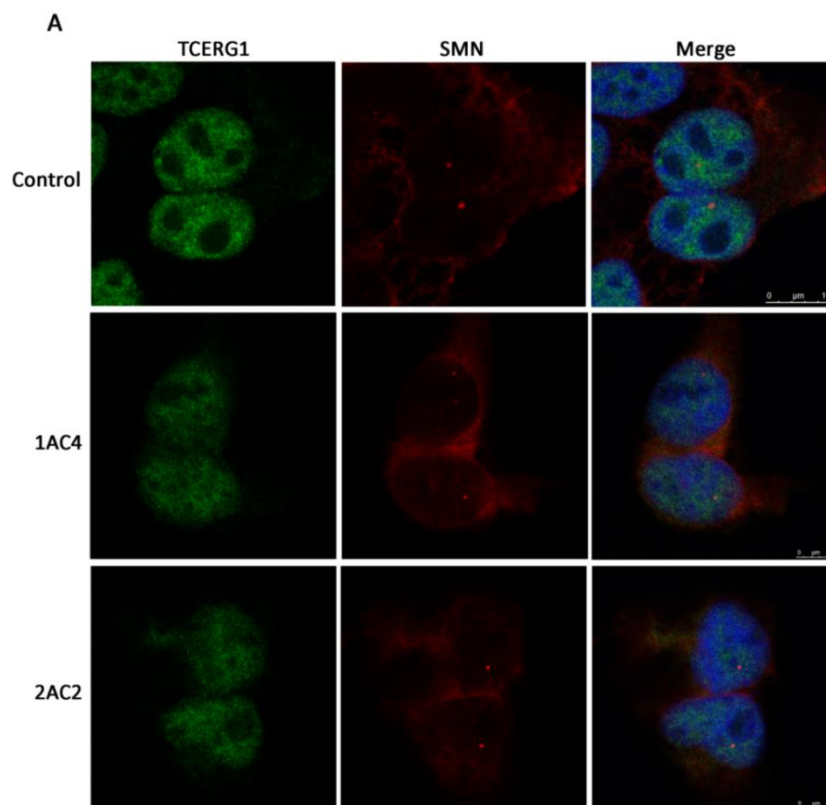
2.4.2. *Gems*

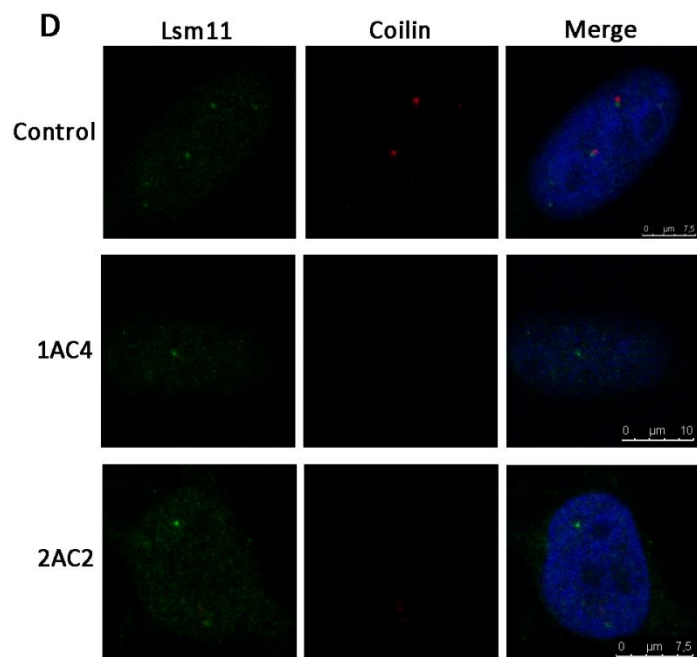
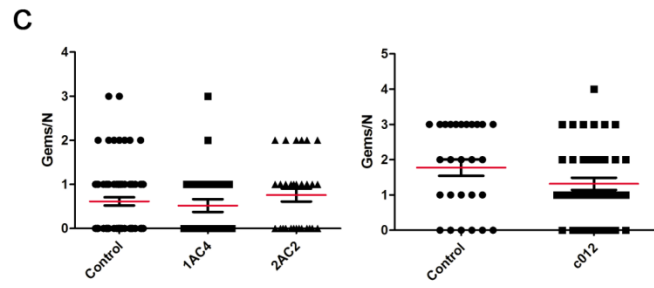
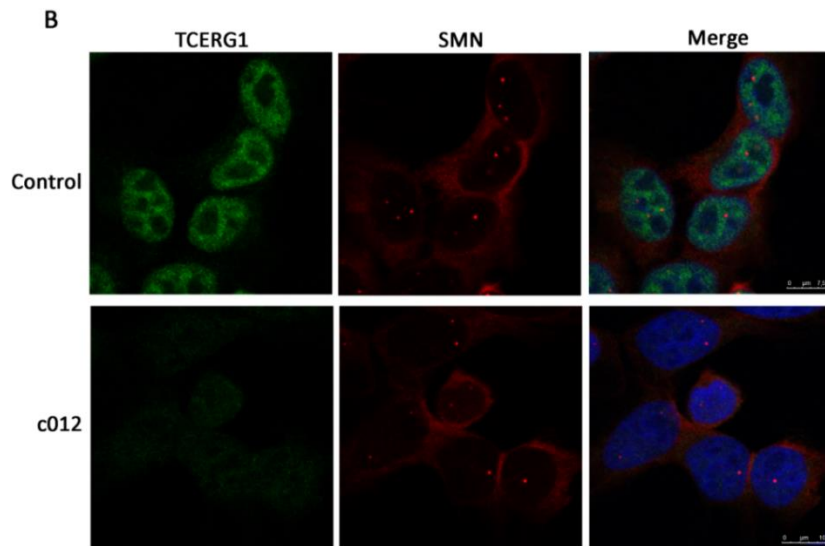
SMN es una proteína que se localiza en los denominados *Gems*, orgánulos estrechamente asociados a los CBs. Al realizar tinciones mediante inmunofluorescencia utilizando un anticuerpo que reconoce la proteína SMN endógena, se comprobó que la ausencia de TCERG1 no afecta a los *Gems*

nucleares. Este estudio se realizó tanto en las líneas CRISPR de HEK293T como en las CRISPR de HAP1 (Fig R15-A-C).

2.4.3. HLB

Los CBs se han relacionado también con los orgánulos nucleares conocidos como *Histone Locus Bodies* (HLBs), el cual se asocia a la cromatina para regular la expresión de los genes de histona. Para determinar si la ausencia de TCERG1 podía afectar a los HLBs como ocurre con los CBs, utilizamos un anticuerpo específico contra Lsm11, que es un componente específico de los HLBs para analizar mediante inmunofluorescencia las líneas HEK293T y HAP1 (Fig R15-D-F). Los resultados obtenidos indicaron que la ausencia de TCERG1 no afectaba la integridad de los HLBs.





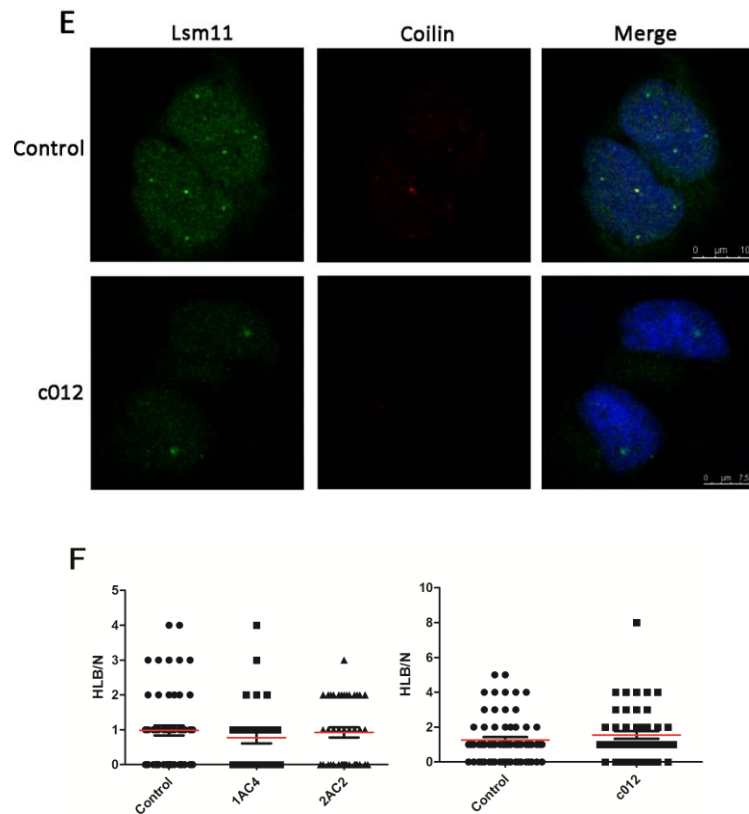


Figura R15. La integridad de *Gems* y HLB no se ve afectada por una disminución de la expresión de TCERG1. Las imágenes muestran el análisis por inmunofluorescencia de células control y silenciadas en HEK293T (A) y HAP1 (B) teñidas con anticuerpos que detectan TCERG1 (verde) y el marcador de los orgánulos nucleares *Gems*, SMN (rojo), así como el solapamiento de ambas señales (Merge). C. Cuantificación de los *Gems* por núcleo (*Gems*/N) en las líneas celulares CRISPR HEK293T (izquierda) y HAP1 (derecha). D, E. Se muestran imágenes de células HEK293T (D) y HAP1 (E) teñidas con anticuerpos que detectan el componente Lsm11 de los HLB (verde) y coilina (rojo), así como la superposición de ambas señales (Merge). F. Cuantificación de los HLB por núcleo (HLB/N) en las líneas celulares CRISPR HEK293T (izquierda) y HAP1 (derecha).

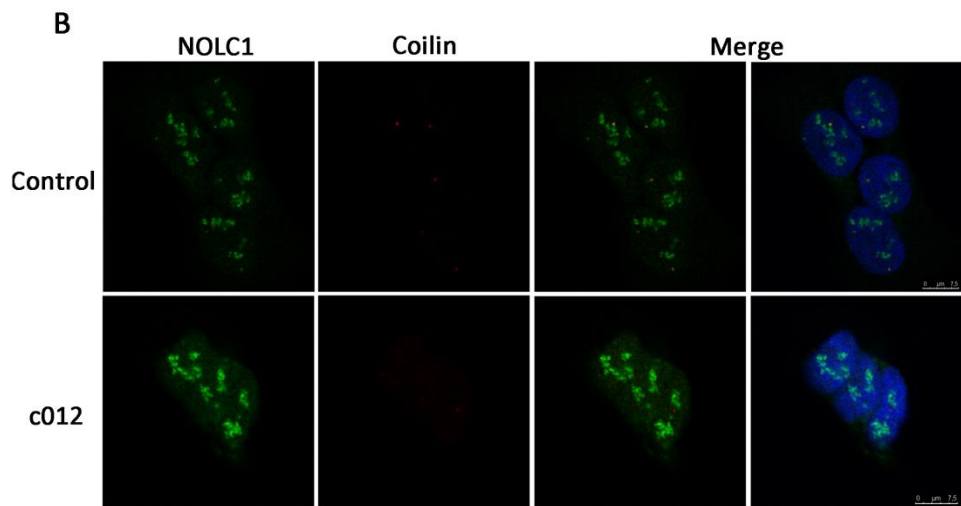
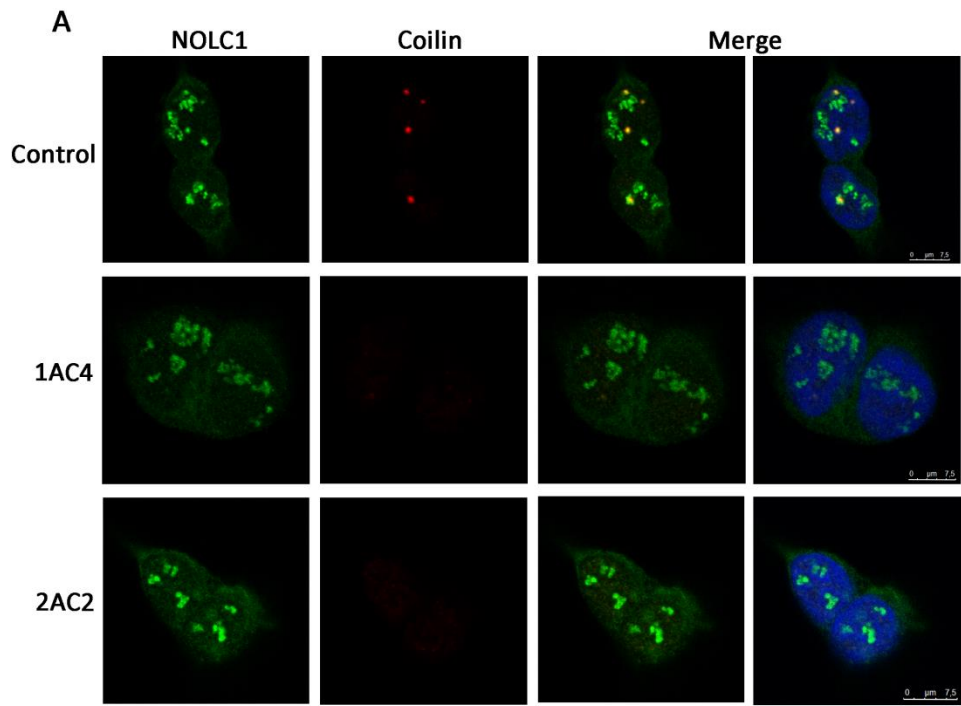
2.4.4. Nucléolo

Como se ha mencionado, NOLC1 se describió como una fosfoproteína que viaja entre nucléolo y el nucleoplasma, donde se acumula en los CBs e interacciona con coilina, sugiriendo que podría actuar como un enlace entre el nucléolo y los CBs. En esta tesis hemos corroborado esos resultados en los experimentos descritos en la Fig. R6. Para comprobar si la ausencia de TCERG1 afectaba la

localización de NOLC1 se realizaron experimentos de inmunofluorescencia usando anticuerpos específicos para NOLC1 y coilina en las líneas HEK293T y HAP1. En las líneas celulares donde la expresión de TCERG1 se encuentra disminuida, NOLC1 se localiza en el nucléolo pero no se acumula en el nucleoplasma, indicando de nuevo una pérdida de la integridad de los CBs debido a la ausencia de TCERG1 sin afectar su localización en el nucleolo (Fig R16 A - B).

Para confirmar que la disminución de la expresión de TCERG1 no afectaba la integridad del nucléolo, se realizó el mismo ensayo de IF con anticuerpos específicos para nucleolina. La nucleolina es una proteína que se encuentra principalmente localizada en el componente fibrilar denso (DFC) y en el componente granular (GC) del nucléolo (Biggiogera et al. 1990). Las imágenes de inmunofluorescencia mostraron que el patrón de localización de esta proteína no cambiaba en condiciones de menor expresión de TCERG1 (Fig. R16 C-D).

Todos estos datos indican que la menor expresión de TCERG1 afecta la integridad de los CBs pero no la de otros orgánulos nucleares.



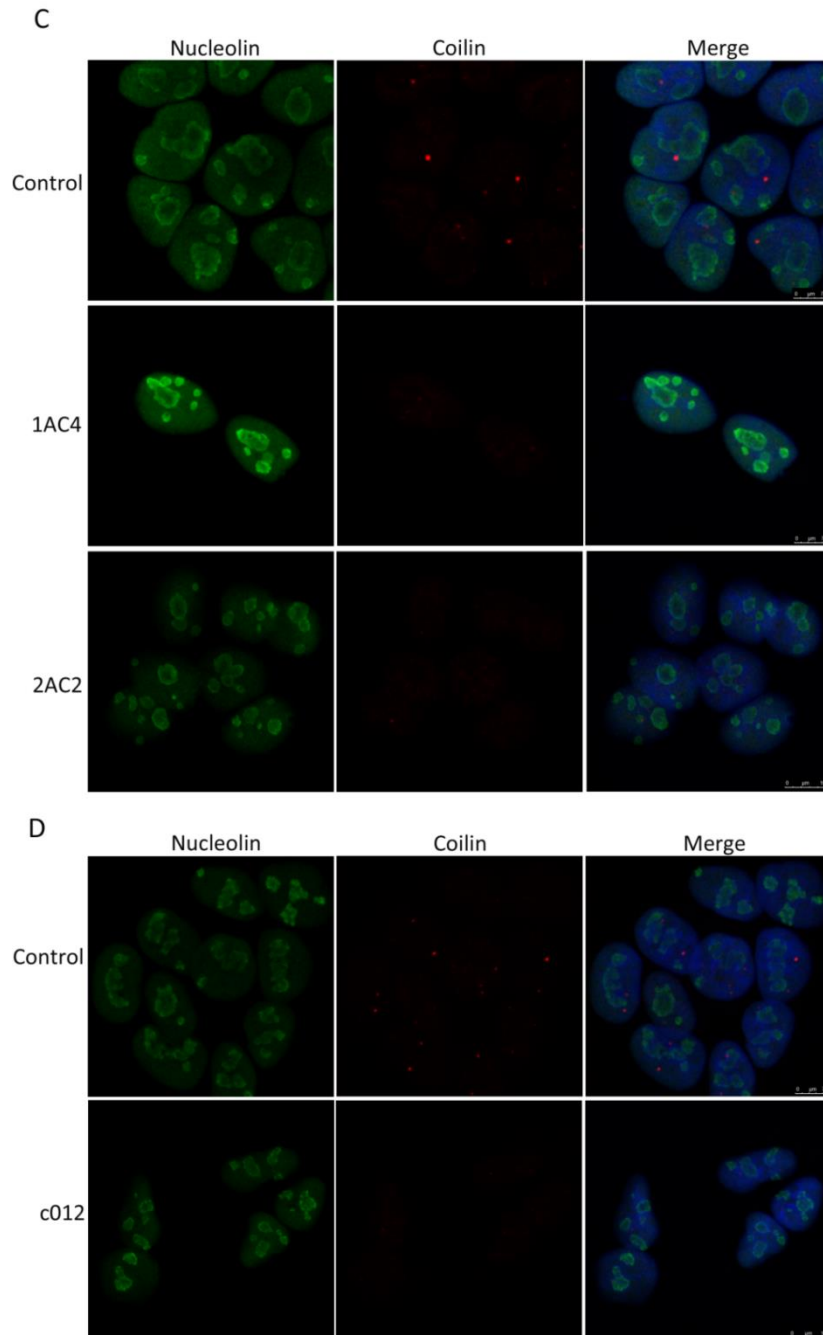


Figura R16. La localización de NOLC1 en los CBs se afecta en ausencia de TCERG1 mientras la localización de NOLC1 en el nucléolo no se ve afectada. El estudio del patrón de distribución de NOLC1 se realizó en experimentos de doble inmunofluorescencia indirecta usando anticuerpos específicos para NOLC1 (verde) y coilina (rojo) usando las líneas CRISPR HEK293T (A) y HAP1 (B). Se muestran además las imágenes superpuestas (Merge) junto con la señal de DAPI (azul). C. Análisis por inmunofluorescencia utilizando anticuerpos frente a nucleolina (verde) coilina (rojo), así como la superposición de ambas señales junto con la señal DAPI (azul)

como marcador de núcleo, en las líneas CRISPR HEK293T. D. El mismo ensayo indicado en C se llevó a cabo usando las líneas CRISPR HAP1.

2.5. La reexpresión de TCERG1 rescata el fenotipo de las líneas silenciadas

Quisimos comprobar si la reexpresión de TCERG1 en las líneas silenciadas podía rescatar el fenotipo observado. Utilizando las líneas celulares CRISPR HEK293T, realizamos una transfección usando cantidades no saturantes (250 ng, ver más abajo) del vector de expresión pEFBOST7/GFP-TCERG1 o del plásmido pEFBOST7 vacío como control y analizamos estas células por microscopía confocal. Las imágenes obtenidas mostraron una restauración de la arquitectura de los CBs en las células de las líneas 1AC4 y 2AC2 donde habíamos re-expresado la proteína TCERG1 con respecto a la línea control (Fig. R17-A). La cuantificación de los resultados indicó un aumento significativo del número de CBs por núcleo y de su tamaño en las células donde habíamos vuelto a expresar TCERG1 con respecto a las células 1AC4 y 2AC2 de la transfección control (Fig. R17-B). Estos resultados apoyan los datos anteriores y sugieren un papel específico de TCERG1 en la formación de los CBs.

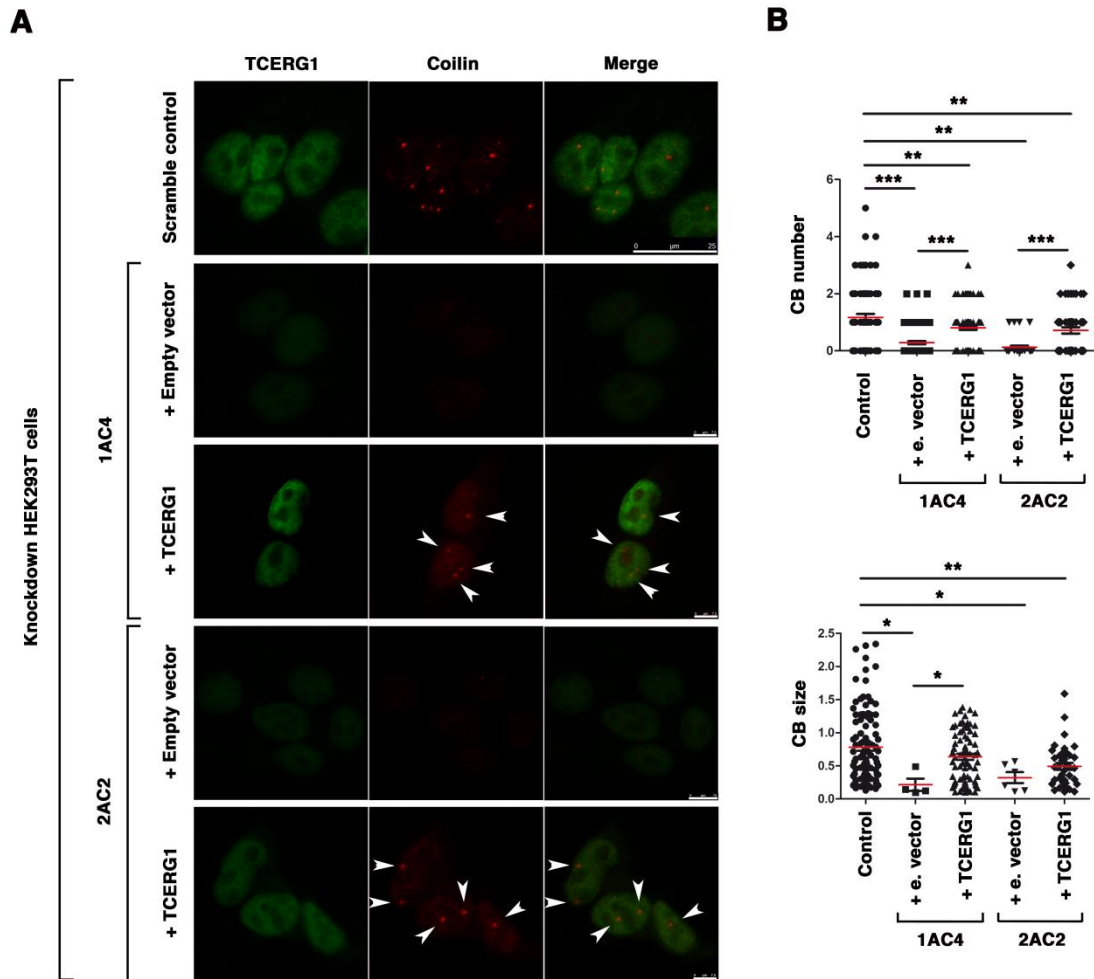


Figura R17. La reexpresión de TCERG1 rescata el fenotipo de las líneas celulares silenciadas. **A.** Las líneas celulares 1AC4 y 2AC2 fueron transitoriamente transfectadas con 250 ng del plásmido pEFBOST7/GFP-TCERG1 o bien con el plásmido vacío pEFBOST7 vacío (*Empty vector*). La línea control fue transfectada con el plásmido vacío. Se obtuvieron imágenes mediante microscopía confocal usando anticuerpos que reconocen TCERG1 (verde) y coilina (rojo). Se muestran además las imágenes superpuestas (Merge). Las flechas señalan ejemplos de CBs en las células transitoriamente transfectadas con el vector de expresión de TCERG1. **B.** Análisis de la cuantificación del número y tamaño de los CBs. Los valores P están representados en las gráficas mediante asteriscos (*, $P < 0,05$; **, $P < 0,01$; ***, $P < 0,005$).

2.6. Efecto de la sobreexpresión saturante de TCERG1 en la integridad de los CBs

Para estudiar si la sobreexpresión saturante de TCERG1 afectaba al patrón de localización de coilina y, por tanto, a la formación de los CBs, realizamos experimentos de inmunofluorescencia utilizando distintas construcciones de TCERG1 unidas a la proteína fluorescente EGFP (*enhanced green fluorescent protein*) junto con el anticuerpo anti-coilin para localizar los CBs. Las imágenes obtenidas mostraron un menor número de CBs en las células donde se sobreexpresó el plásmido que codifica el gen *TCERG1* WT (Figura R18, EGFP-TCERG1-WT). Para determinar la región de TCERG1 responsable de este efecto, se realizaron transfecciones transitorias con construcciones que expresan los extremos amino y carboxiterminal de TCERG1 de forma independiente y fusionadas a EGFP. Sólo la construcción que expresaba el extremo carboxiterminal de TCERG1, donde se encuentran los dominios FF, afectó la integridad de los CBs (Figura R18; EGFP-TCERG1-Nt, EGFP-TCERG1-Ct). A continuación, quisimos conocer los dominios responsables de este fenotipo. Sobreexpresando distintas construcciones truncadas de TCERG1, observamos que los dominios FF4 y FF5 eran los responsables de que la localización de coilina en los CBs estuviera afectada. La sobreexpresión de una proteína de fusión que contenía el extremo amino-terminal de TCERG1, que ya habíamos observado que no afectaba la localización de coilina, junto con los dominios FF4 y FF5 produjo un fenotipo similar al observado con la proteína TCERG1-WT con la deslocalización de coilina en las células sobre-expresadas (Fig. R18; EGFP-TCERG1(N)t-FF4/FF5). Corroborando los anteriores resultados, la sobreexpresión de una proteína de fusión conteniendo la secuencia de TCERG1 completa excepto los dominios FF4 y FF5 produjo un fenotipo similar al que se da en las células no transfectadas (Fig. R18; EGFP-TCERG1- Δ FF4/FF5). Por ello se concluye que los dominios FF4 y FF5 de TCERG1 son los responsables del fenotipo de los CBs observado, tras la sobreexpresión de la proteína. Dada la relevancia de

los residuos de fenilalanina de los dominios FF4/FF5 en la interacción entre TCERG1 y NOLC1 (Fig. R2), quisimos conocer la implicación de estos aminoácidos en la integridad de los CBs en los experimentos de sobreexpresión. La sobreexpresión de la proteína de fusión con mutaciones en estos residuos no modificó la localización de coilina en el nucleoplasma y por tanto la integridad de los CBs (Figura R18, EGFP-TCERG1(Nt)-FF4/FF5 [F903, 961A]). Por tanto, las fenilalaninas 903 y 961 presentes en los dominios FF4 y FF5 de TCERG1 son importantes para la integridad de los CBs en condiciones de sobreexpresión saturante de TCERG1.

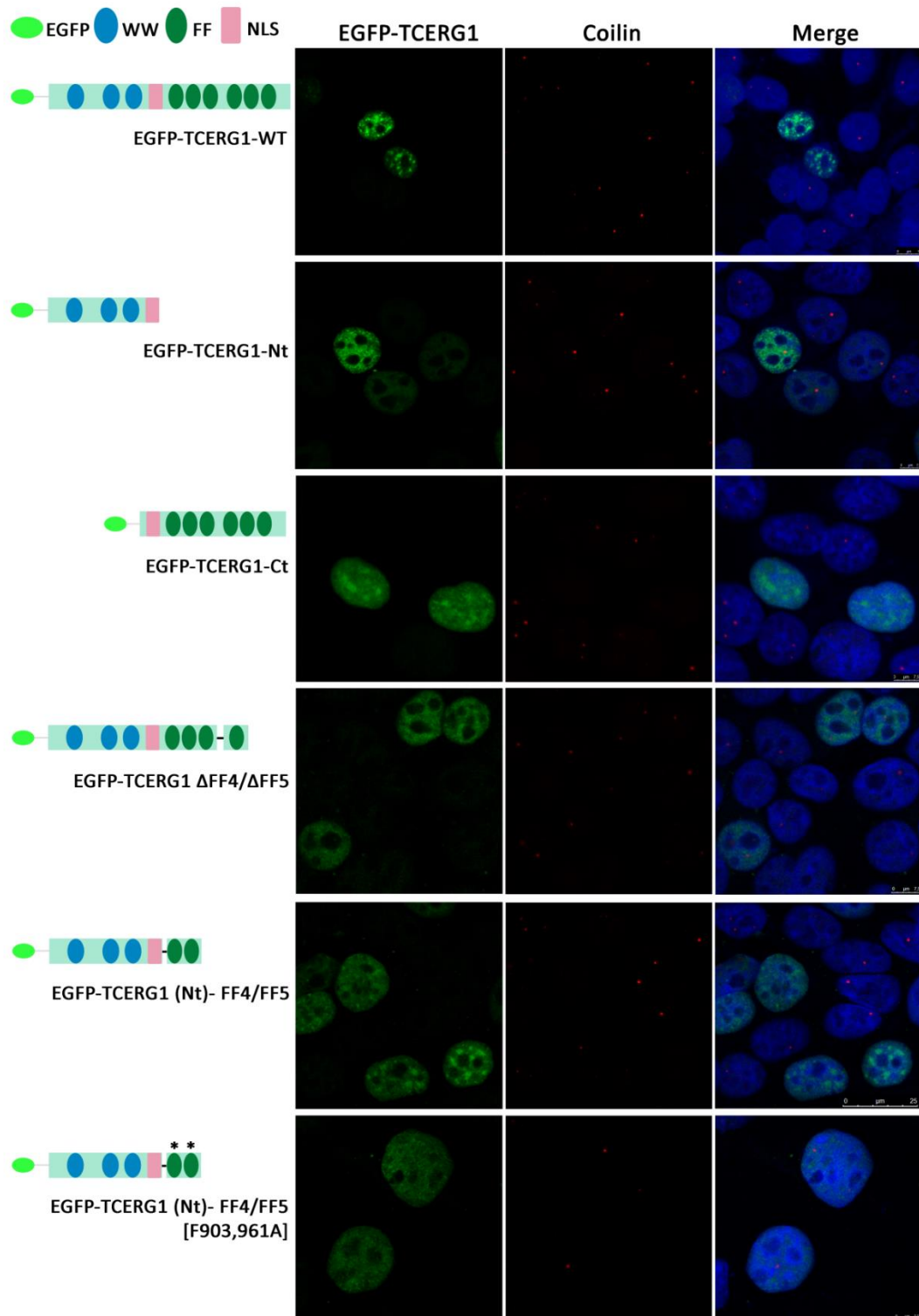


Figura R18. La sobreexpresión de TCERG1 afecta la integridad de los CBs. Las imágenes muestran células HEK293T expresando construcciones indicadas de EGFP-TCERG1 (verde) y la localización de coilina usando anticuerpos específicos (rojo). El diagrama estructural de cada construcción se representa a la izquierda de cada panel. En los diagramas se indican la proteína EGFP, los dominios WW, la señal de localización nuclear (NLS) y los dominios FF; los asteriscos representan los residuos mutados en la construcción. En cada panel se muestran las imágenes individuales (EGFP-TCERG1 y Coilin) y superpuestas (Merge).

A continuación quisimos conocer si el fenotipo observado tras la sobreexpresión de TCERG1 era específico de esta proteína. Para ello, realizamos similares experimentos de sobreexpresión con el factor de *splicing* SF1 en células HEK293T. La sobreexpresión de SF1 no afectó la localización de coilina (Fig.R19-A). El mismo resultado se observó al transfectar el plásmido pEFBOST7-GFP sobreexpresando la proteína de fusión T7-EGFP (Fig.R19-B). La sobreexpresión de TCERG1 desde un vector diferente al usado en los experimentos de la Fig. R16 (pEFBOST7-TCERG1) afectó de nuevo la localización de coilina, confirmando el efecto específico de la sobreexpresión saturante de TCERG1 en la integridad de los CBs (Fig. R19-C).

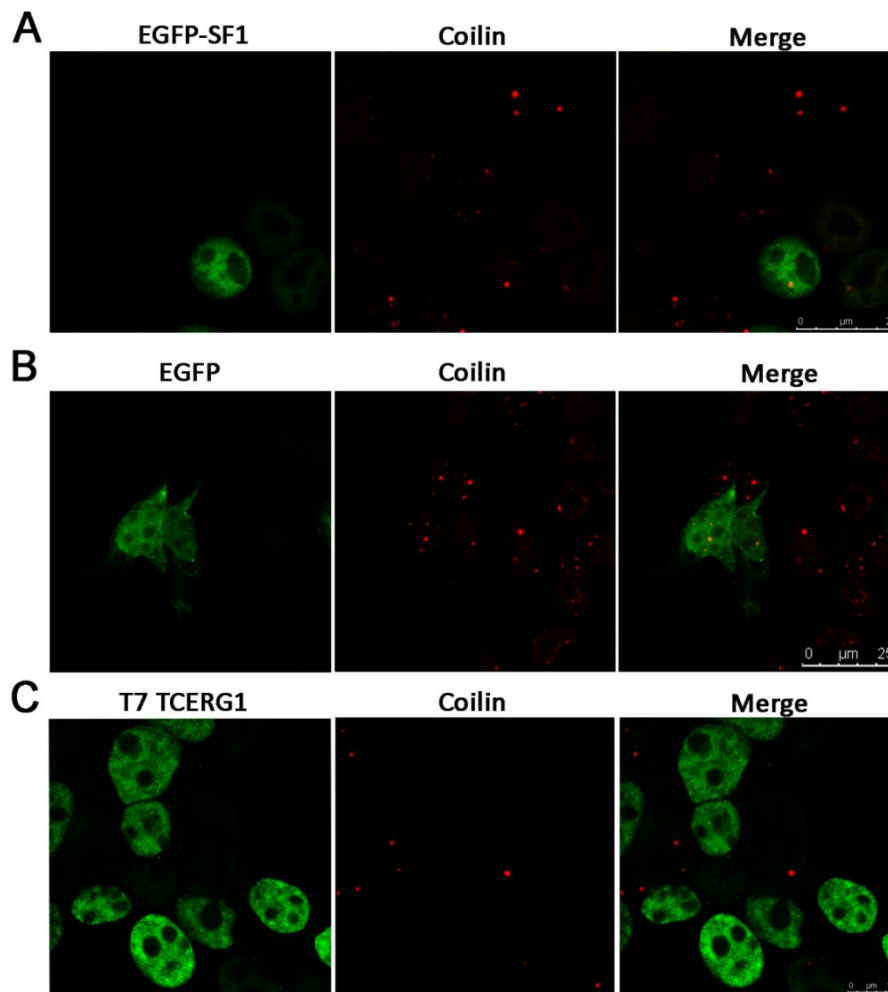


Figura R19. El efecto de la sobreexpresión de TCERG1 en la integridad de los CBs es **específico**. La sobreexpresión de la proteína de fusión de SF1 con EGFP (EGFP-SF1) (A) o la sobreexpresión de EGFP (T7-EGFP) (B) no afectó el patrón de localización nuclear de coilina en células HEK293T. C. La sobreexpresión de TCERG1 usando el vector pEFBOST7-TCERG1 afectó la localización de coilina en células HEK293T. Las proteínas TCERG1 y coilina se detectaron usando anticuerpos específicos para el epítipo T7 (verde) y para coilina (rojo), respectivamente.

2.7. La sobreexpresión de TCERG1 no afecta la expresión de coilina.

A la vista de estos resultados, quisimos comprobar si los niveles de coilina estaban siendo afectados por una mayor expresión de TCERG1. Para ello, células HEK293T se transfectaron con el plásmido de expresión pEFBOST7-TCERG1 o con

el plásmido pEFBOST7 como control (Mock) y los niveles de proteína se analizaron mediante WB usando anticuerpos específicos. La sobre-expresión de TCERG1 no afectó los niveles de expresión de esta proteína (Fig. R20-A). Estas mismas células transfectadas se utilizaron para realizar una extracción de RNA y comprobar por RT-qPCR que los niveles de mRNA de coilina tampoco se afectaban por la sobreexpresión de TCERG1 (Fig. R20-B).

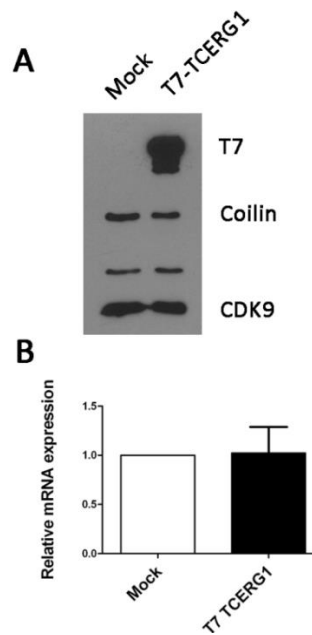


Figura R20. La sobreexpresión de TCERG1 no afecta la expresión de coilina. A. Análisis de extractos de células HEK293T transfectadas con el plásmido pEFBOST7-TCERG1 o el plásmido pEFBOST7 (Mock) mediante *western blot* usando anticuerpos específicos para el epítipo T7 y la proteína coilina. El anticuerpo anti-CDK9 se usó como control. **B.** Análisis mediante RT-qPCR de los niveles de mRNA de coilina utilizando GAPDH como control interno en células control (Mock) y transfectadas con el plásmido de expresión pEFBOST7-TCERG1.

2.8. La sobreexpresión de TCERG1 no afecta la localización de otros orgánulos nucleares

Vistos los resultados observados sobre los CBs al sobre-expresar TCERG1 quisimos comprobar que, al igual que ocurre en el caso del silenciamiento de esta

proteína, la sobreexpresión de TCERG1 no afecta a otros orgánulos nucleares. Con este fin realizamos experimentos de inmunofluorescencia indirecta utilizando los anticuerpos que reconocen a las proteínas marcadoras de *speckles*, *gems* y HLBs en condiciones de sobreexpresión de TCERG1. Ninguno de estos cuerpos nucleares fue alterado por la sobreexpresión de TCERG1 (Fig. 21A-C). Por tanto, se concluye que la sobreexpresión de TCERG1 no afecta a distintos orgánulos nucleares muy relacionados entre ellos como son los *speckles*, *gems* o HLBs y que solo afecta al patrón de localización de coilina, y por tanto a la integridad de los CBs.

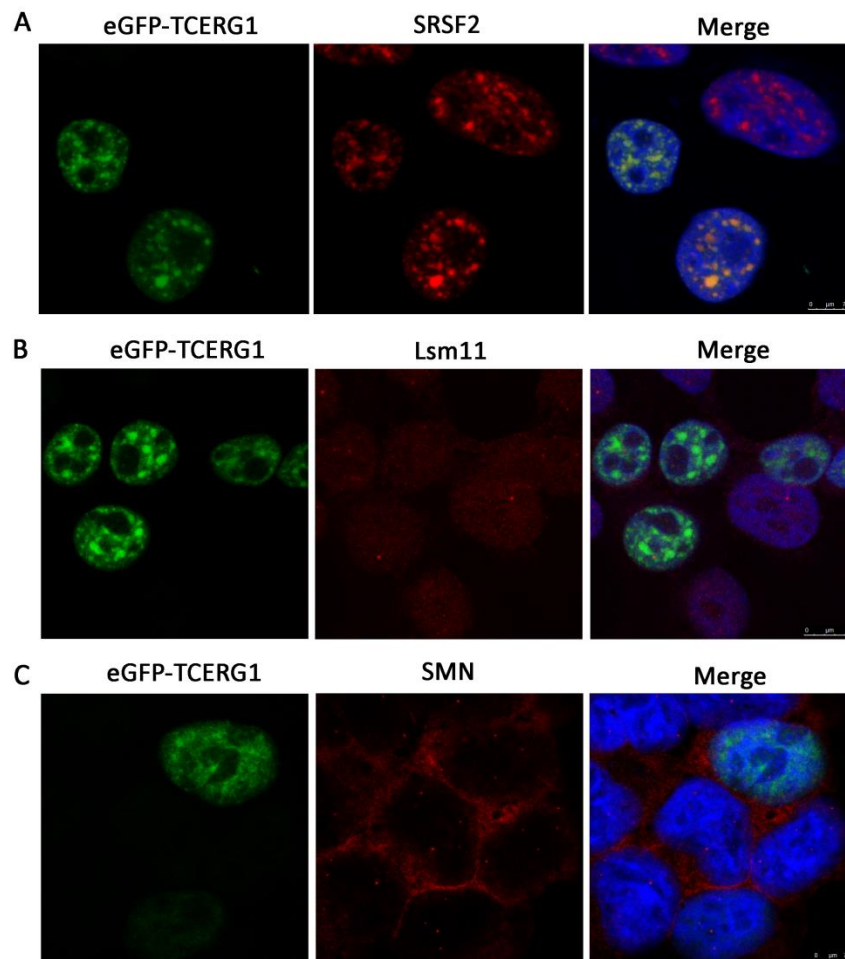


Figura R21. La sobreexpresión de TCERG1 no afecta la integridad de *speckles*, HLB ni *gems*. Análisis del efecto de la sobreexpresión de TCERG1 (eGFP-TCERG1, verde) sobre distintos orgánulos nucleares (rojo) mediante microscopía confocal usando anticuerpos que reconocen SRSF2 (A), Lsm11 (B) y SMN (C), que marcan *speckles*, HLBs y *gems*, respectivamente.

Como ya se ha comentado anteriormente, NOLC1 es una fosfoproteína que se concentra en el nucléolo y en los CBs. Una menor expresión de TCERG1 afecta a su patrón de localización en el nucleoplasma pero no a su localización en el nucléolo (Fig. 16A-B). Para estudiar si la sobreexpresión de TCERG1 afecta la distribución de NOLC1 en el núcleo, células HEK293T se transfectaron con la construcción eGFP-TCERG1 y se estudió la localización de coilina y NOLC1 por inmunofluorescencia. Como era de esperar, NOLC1 se localizó en el nucleolo y en los CBs, detectados mediante coilina, en las células no transfectadas. En las células transfectadas donde era visible la sobreexpresión de TCERG1, sin embargo, NOLC1 sólo se encontró presente en el nucleolo pero no en el nucleoplasma (Fig. R22). Estos datos son consistentes con los anteriores resultados que sugieren la pérdida de integridad de los CBs en condiciones de sobreexpresión de TCERG1.

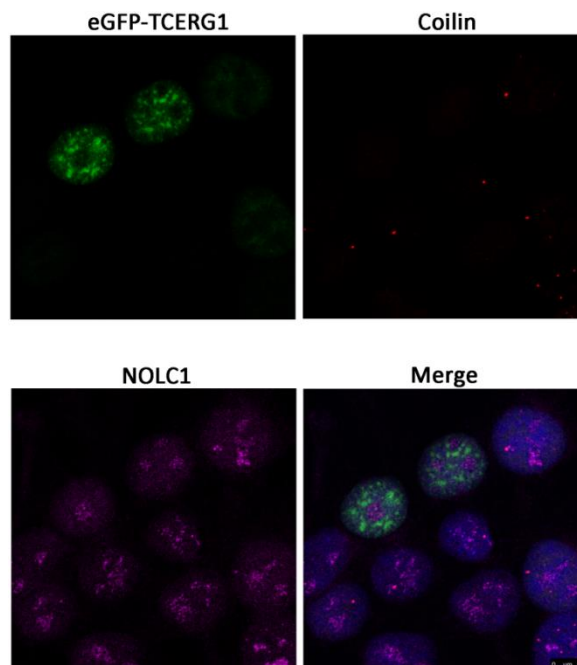


Figura R22. La sobreexpresión de TCERG1 afecta la localización de NOLC1 en los CBs. Las imágenes muestran células HEK293T expresando la construcción eGFP-TCERG1 (verde) y

marcadas con anticuerpos que reconocen coilina (rojo) y NOLC1 (magenta), así como la superposición de las tres señales (Merge) junto con el marcaje en azul por DAPI.

2.9. Formación de homocomplejos de TCERG1

Hemos observado que la ausencia de TCERG1 afecta la integridad de los CBs, sugiriendo que esta proteína es importante para el mantenimiento de estos orgánulos nucleares. El hallazgo de que TCERG1 exógeno altera la localización de coilina sugiere un efecto dominante negativo de TCERG1 sobreexpresado que podría ser causado por la autointeracción de TCERG1. Trabajos anteriores han demostrado este fenómeno para coilina, donde la proteína sobreexpresada altera los CBs a través de un efecto dominante negativo entre la coilina exógena y endógena (Hebert and Matera 2000). Para investigar si la proteína TCERG1 se asocia consigo misma, se utilizaron dos vectores de expresión que contenían la secuencia de TCERG1 y dos epítomos diferentes: pEFBOST7-TCERG1 que contiene el epítomo T7 y pcDNA3/EGFP/TCERG1 que contiene el epítomo GFP. Como controles se utilizaron los vectores vacíos pEFBOST7 y pcDNA3/EGFP. Células HEK293T se cotransfectaron transitoriamente con esas construcciones y los extractos inmunoprecipitados con anti-T7 y anti-GFP. El análisis mediante *western blot* mostró que T7-TCERG1 coprecipita EGFP-TCERG1 y viceversa (Fig. R23). Estos resultados indican que TCERG1 exógeno autointeracciona *in vivo* y sugieren la hipótesis de que la sobreexpresión de TCERG1 puede “secuestrar” a TCERG1 endógeno y causar la pérdida de integridad de los CBs.

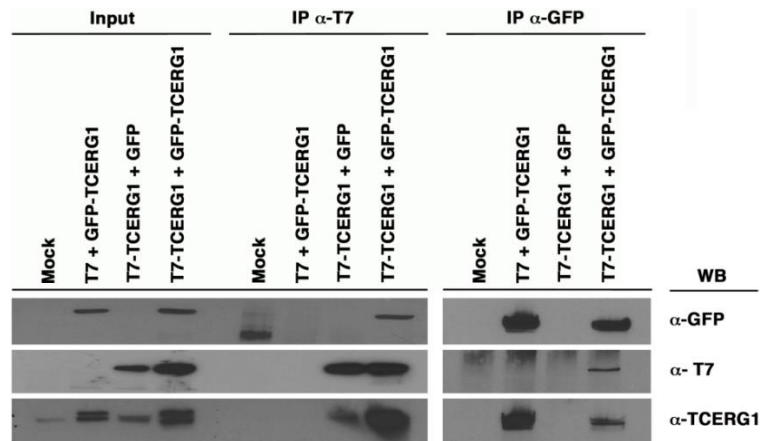


Figura R23. Autointeracción de TCERG1. Se prepararon extractos celulares de células HEK293T transfectadas transitoriamente con los plásmidos de expresión pEFBOST7-TCERG1 o vacío (T7) o con los plásmidos de expresión pcDNA3/EGFP/TCERG1 o vacío (eGFP). Una fracción de los extractos se guardó (Input) y el resto se inmunoprecipitó con anticuerpos específicos anti-T7 y anti-GFP. Las proteínas inmunoprecipitadas se separaron por SDS-PAGE y se analizaron por WB utilizando anticuerpos específicos para GFP, T7 y TCERG1.

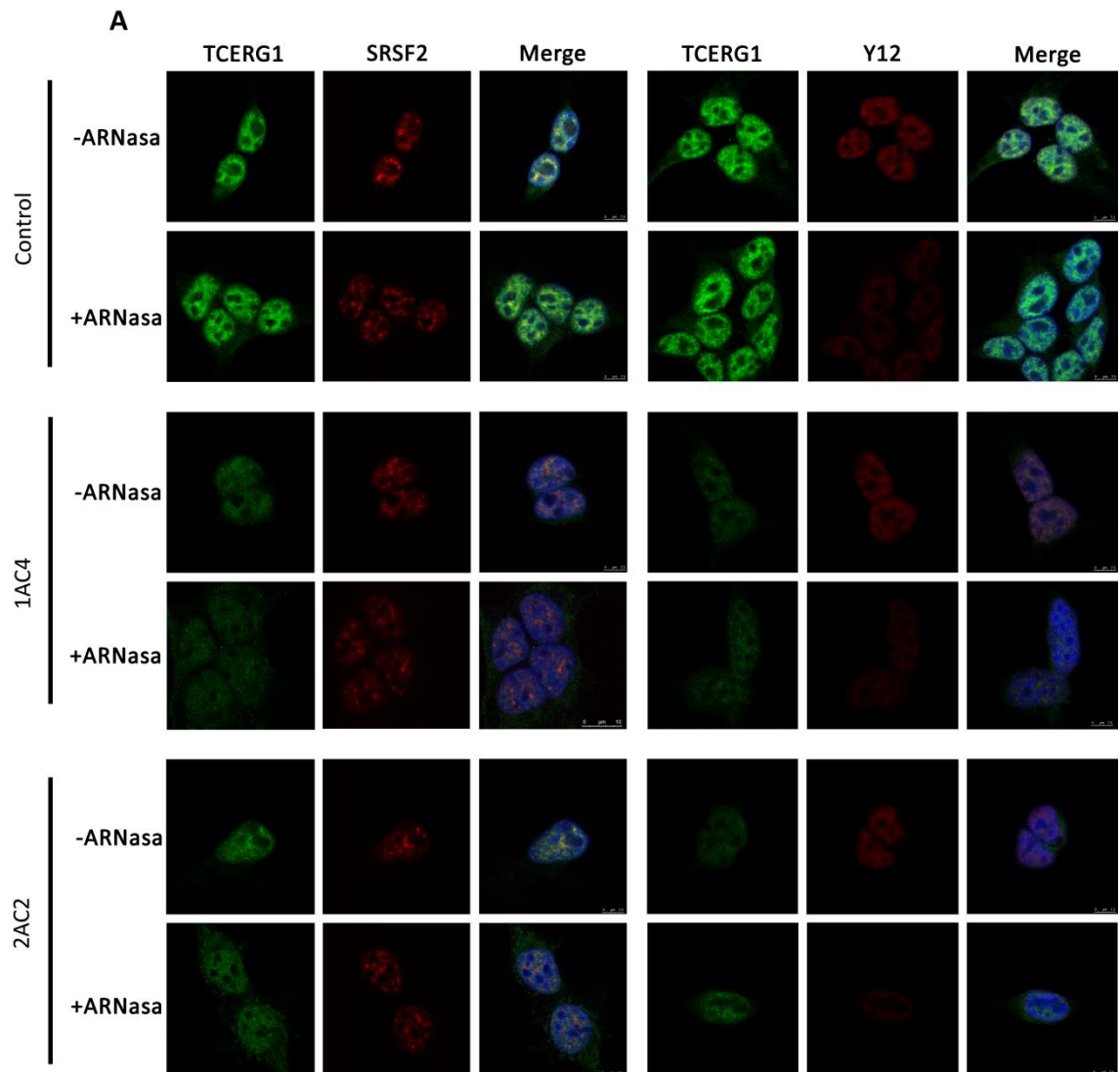
2.10. El efecto de TCERG1 sobre los CBs no está mediado por RNA

Como ya se mencionó en la Introducción, los CBs son estructuras donde se ensamblan y maduran los snRNPs y donde se concentran distintos tipos de RNAs. Por otra parte, la interacción entre TCERG1 y NOLC1 parece estar mediada por RNA (Fig. R5). Partiendo de esta base, decidimos estudiar la importancia del RNA en el fenotipo observado por la inhibición/sobreexpresión de TCERG1. Para ello se usaron dos controles: el factor SRSF2, cuya localización es resistente al tratamiento con RNAsa y las ribonucleoproteínas Sm, cuya localización es sensible al tratamiento con RNAsa.

En primer lugar, realizamos el tratamiento con RNAsa en las líneas celulares CRISPR HEK293T. En estas condiciones, el patrón de localización de SRSF2 y TCERG1, tanto en la línea control como en las líneas silenciadas, no cambió

respecto a los controles sin tratar mientras que la localización de las ribonucleoproteínas Sm cambió para distribuirse de forma difusa por el nucleoplasma (Fig. R24-A). Asimismo, el patrón de localización de coilina no se modificó con el tratamiento con RNAsa en las células control, ni cambió el efecto de la disgregación de los CBs en las líneas 1AC4 y 2AC2 (Fig. R24-B, paneles izquierdos). Por lo tanto, concluimos que el efecto del silenciamiento de TCERG1 producido en la localización de coilina no está mediado por RNA. Curiosamente, El patrón de localización de NOLC1 en las células control sí se modificó con el tratamiento de RNAsa con respecto a las células no tratadas, mostrando una señal más difusa y concentrándose en el citoplasma. El patrón de localización de NOLC1 con el tratamiento de RNAsa en las líneas 1AC4 y 2AC2 mostró un efecto aún más visible, ya que se perdía incluso la tinción de NOLC 1 en el nucléolo (Fig. R24-B, derecha).

En segundo lugar, realizamos el tratamiento con RNAsa en células HEK293T transfectadas de forma transitoria con el vector de expresión de TCERG1. Al igual que lo observado en el experimento anterior, el patrón de localización de TCERG1 y SRSF2 no cambió con respecto a los controles sin tratar, mientras que las ribonucleoproteínas Sm se localizaron de forma difusa en todo el nucleoplasma (Fig. R24-C, D). De la misma manera que ocurrió con las líneas silenciadas, observamos que el RNA no parece intervenir en el efecto de disociación de los CBs en las células donde se sobreexpresó TCERG1 (Fig. R24-E). Igualmente, observamos que el cambio en la distribución de NOLC1 tras el tratamiento con RNAsa no se vió afectado por la sobreexpresión de TCERG1.



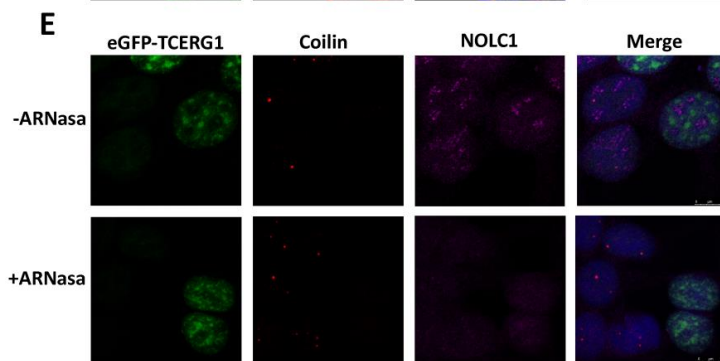
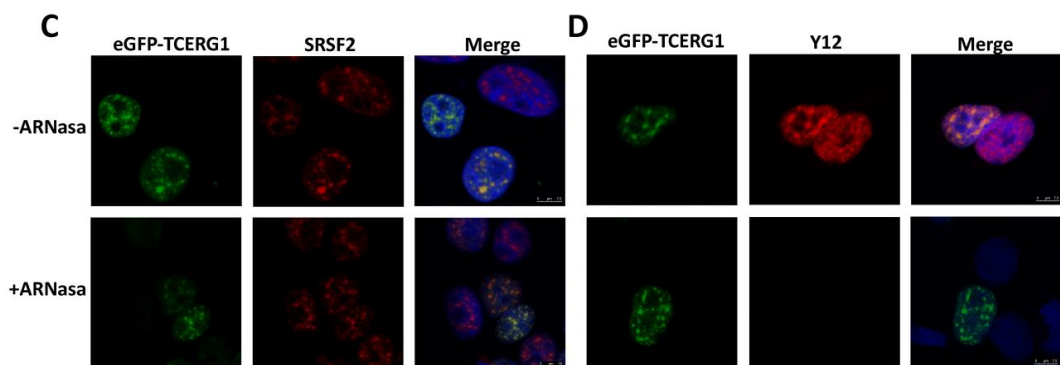
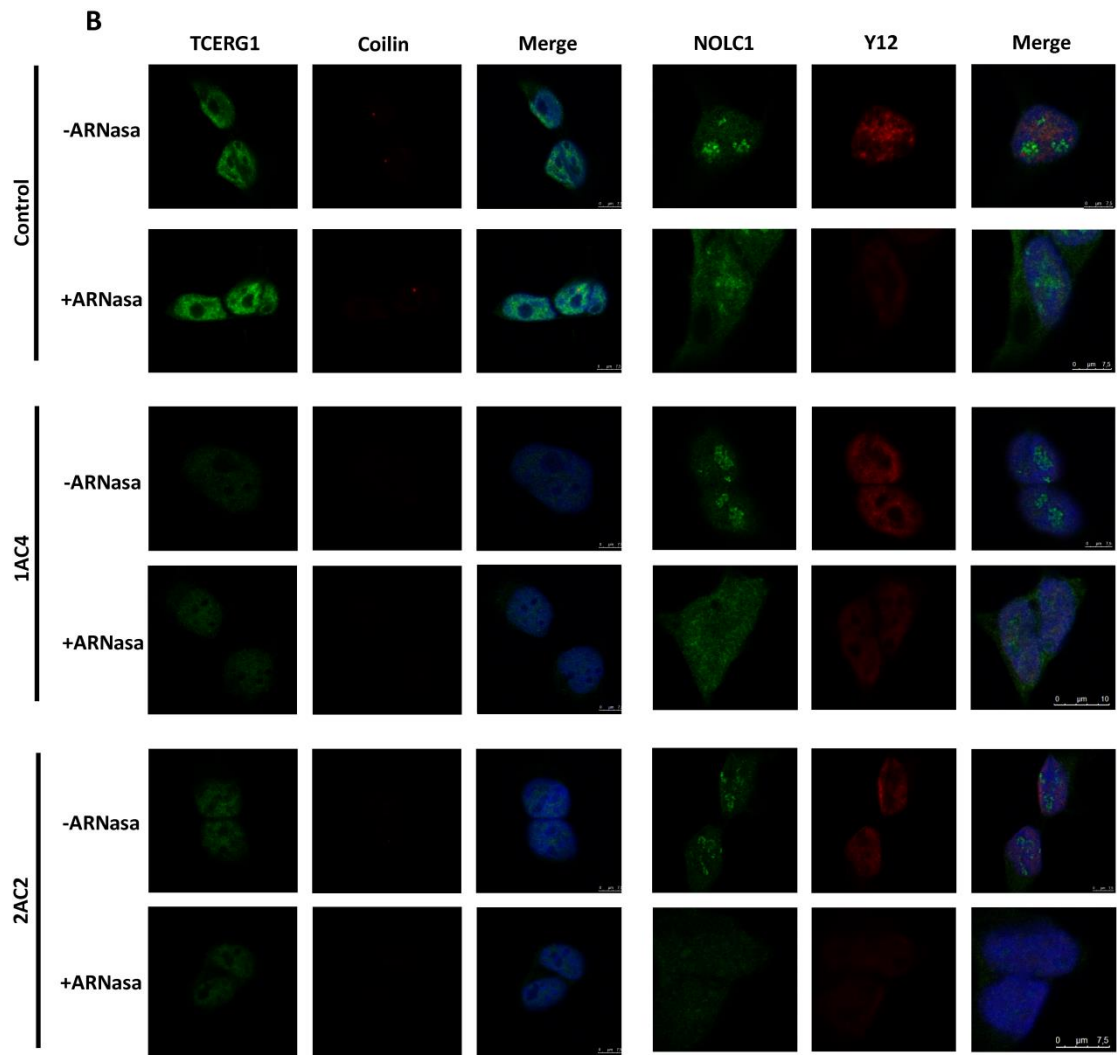


Figura R24. La disociación de los CBs producida por la inhibición/sobreexpresión de TCERG1 no depende de RNA. A. Las células CRISPR HEK293T fueron tratadas con RNAsa A (0,1 mg/ml, 2h a 37°C). La proteína TCERG1 se visualizó utilizando un anticuerpo específico (verde). La distribución de SRSF2 (izquierda) y Sm (derecha) (rojo) se utilizó como control negativo y positivo del tratamiento, respectivamente. Las imágenes muestran las tinciones individuales y la superposición de ambas (Merge) sin el tratamiento (paneles superiores) y con el tratamiento con RNAsa A (paneles inferiores). Se muestran las imágenes superpuestas junto con la señal en azul por DAPI (Merge). B. Imágenes realizando el mismo tratamiento con RNAsa y marcando con un anticuerpo específico que reconoce coilina (rojo). Se muestran las imágenes superpuestas junto con la señal en azul por DAPI. C, D. Las células HEK293T se transfectaron con el vector de expresión de TCERG1 (eGFP-TCERG1) y se trataron con RNAsa A (0,1 mg/ml, 2h a 37°C). La distribución de SRSF2 (C) y Sm (D) se verificó con anticuerpos anti-SRSF2 y anti-Sm (Y21) (rojo). Se muestran las imágenes superpuestas junto con la señal en azul por DAPI (Merge). E. Las imágenes muestran la sobreexpresión de eGFP-TCERG1 (verde) y el marcaje con anticuerpos específicos que reconocen las proteínas coilina (rojo) y NOLC1 (magenta). Se muestran las imágenes superpuestas junto con la señal en azul por DAPI (Merge).

3. Efecto del silenciamiento de NOLC1 y coilina en la integridad de los CBs

3.1. El silenciamiento de NOLC1 afecta a los CBs

Quisimos comprobar el efecto de una menor expresión de coilina y de NOLC1. Para ello, sembramos células HEK293T sobre portaobjetos y las transfectamos con un siRNA contra coilina, contra NOLC1 o bien contra GFP a modo de control. Bajo estas condiciones, analizamos mediante inmunofluorescencia los cambios producidos tanto en coilina como en NOLC1. Al utilizar el siRNA específico para coilina (siCoilin) observamos que NOLC1 solo se acumulaba en el nucléolo y no se observaban los acúmulos en el nucleoplasma

coincidentes con los CBs (Fig. R25-A, siCoilin). Al silenciar NOLC1, se observó un número disminuido de CBs y la presencia de coilina en numerosos microgránulos por el nucleoplasma (Fig. R25-A, siNOLC1). El efecto de los siRNAs utilizados en la expresión de las proteínas se verificó mediante WB (Fig. R25-B).

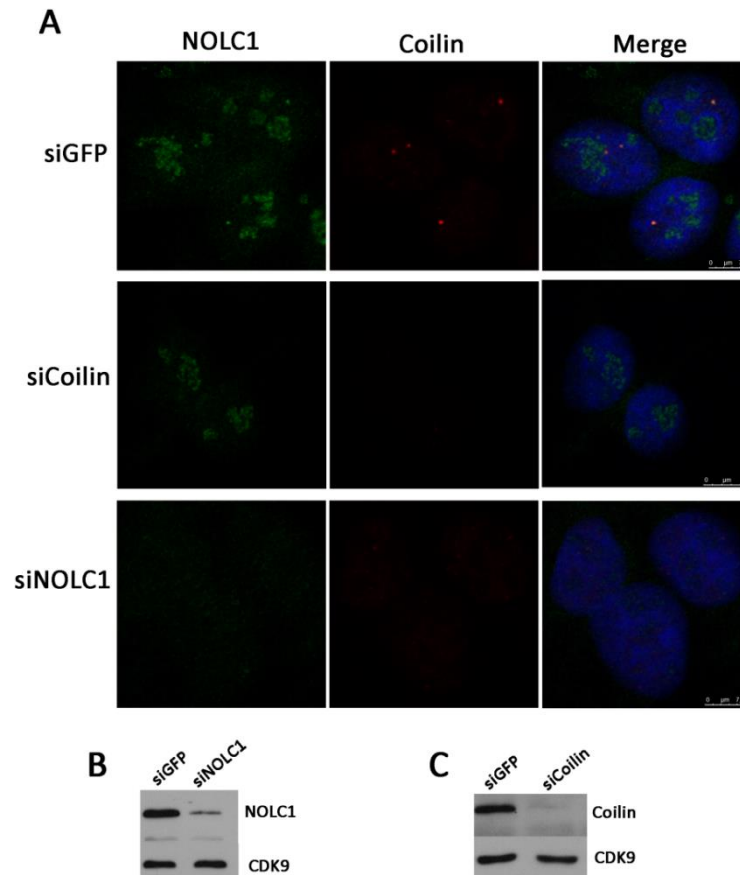


Figura R25. Efecto del silenciamiento de coilina y NOLC1. A. Las imágenes muestran el análisis por inmunofluorescencia de células HEK293T transfectadas con siRNAs específicos para GFP (siGFP), coilina (siCoilin) y NOLC1 (siNOLC1) usando anticuerpos que detectan NOLC1 (verde) y coilina (rojo) así como el solapamiento de ambas señales junto con la señal en azul por DAPI (Merge) B,C. Una fracción de los lisados celulares fue analizada por *western blot* utilizando anticuerpos específicos de NOLC1 y coilina para comprobar el grado de silenciamiento. El anticuerpo anti-CDK9 se usó como control.

3.2. El silenciamiento de NOLC1 y coilina no afecta la integridad de otros orgánulos nucleares

A continuación quisimos comprobar si el silenciamiento de coilina o de NOLC1 podía afectar a otros orgánulos nucleares, en concreto los *gems* y el nucléolo. Ninguno de estos orgánulos se vio afectado por el silenciamiento de coilina o de NOLC1 (Fig. R26-A).

Dado que habíamos observado cambios en la localización de NOLC1 y coilina debido a la disminución de la expresión de TCERG1, nos pareció interesante analizar si TCERG1 estaba asimismo afectado por una menor expresión de coilina y NOLC1. Para ello, se realizó un análisis por IF utilizando anticuerpos específicos dirigidos contra las proteínas endógenas TCERG1 (verde), coilina y SRSF2 (rojo). Los resultados mostraron que la localización de TCERG1 no se vio afectada por el silenciamiento de esas proteínas (Fig. R26-C,D). Estos análisis además, indicaron que una menor expresión de coilina o NOLC1 no afectaba la localización de los *speckles* nucleares, detectada usando el marcador SRSF2. Con estos resultados concluimos que una menor expresión de coilina y NOLC1 afecta específicamente la integridad de los CBs, pero no afecta la localización de la proteína TCERG1.

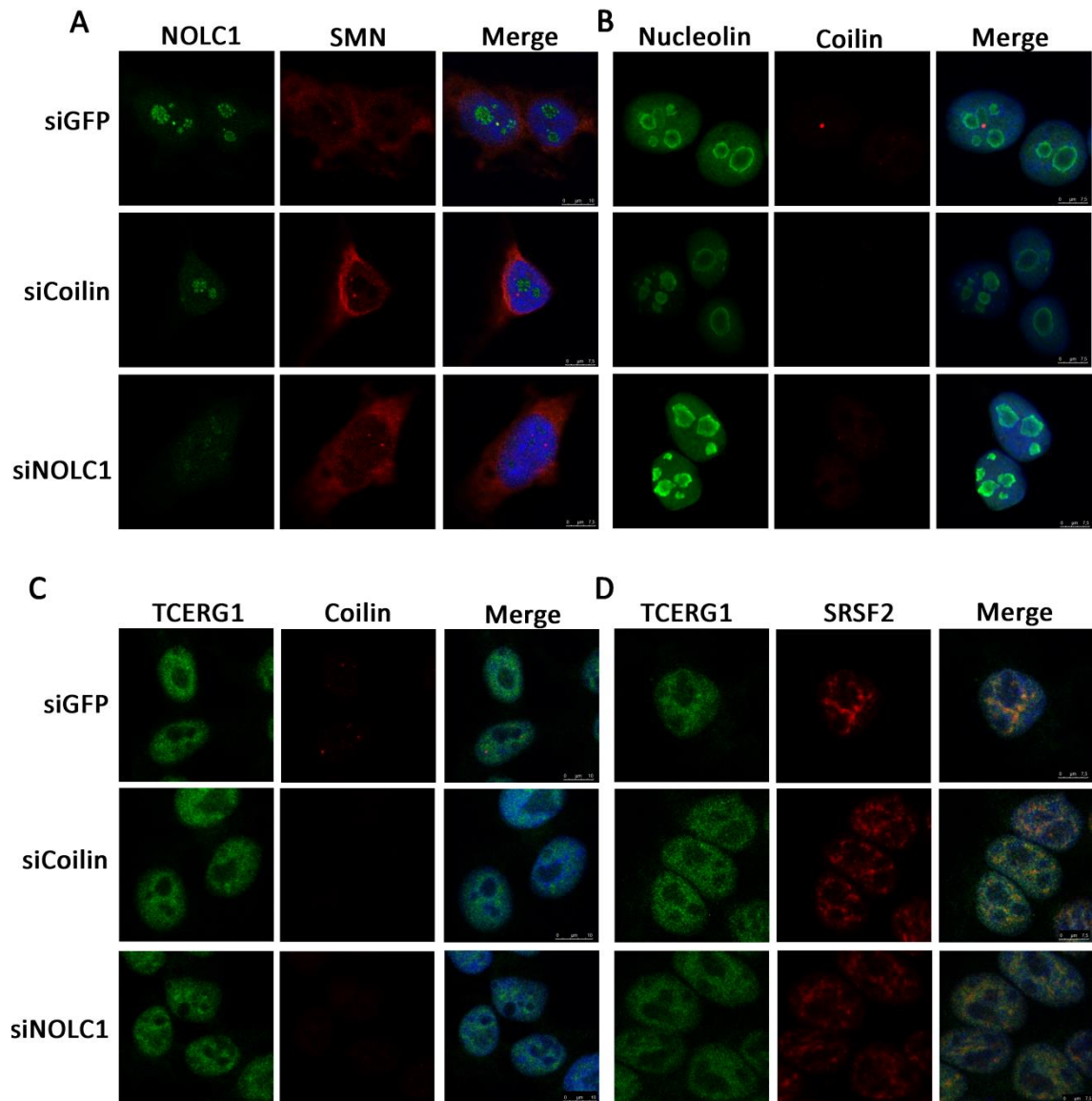


Figura R26. El silenciamiento de coilina o NOLC1 no afecta la integridad de *gems* y nucléolo ni la localización de TCERG1. **A.** Las imágenes muestra el análisis por IF de células transfectadas con siRNAs específicos para GFP (siGFP), coilina (siCoilin) y NOLC1 (siNOLC1) usando anticuerpos que reconocen las proteínas endógenas NOLC1 (verde) y el marcador de *gems* SMN (rojo) así como la superposición de ambas señales (Merge). **B.** En la figura se muestra el doble marcaje realizado con anticuerpos frente a nucleolina (verde) y coilina (rojo), así como la superposición de ambas señales (Merge). **C, D.** El mismo experimento descrito en A y B se analizó por microscopía usando anticuerpos específicos para TCERG1 (verde), coilina (rojo, C) y SRSF2 (rojo, D). Se muestran además, las imágenes de la superposición de las señales verde y rojo junto con la señal en azul por DAPI (Merge) en ambos casos.

4. Papel de TCERG1 en la biosíntesis de snRNPs

4.1. TCERG1 interacciona con distintos componentes relacionados con la biogénesis de los snRNPs.

En los CBs tiene lugar el proceso de formación y maduración de los snRNPs. En este proceso intervienen un gran número de factores, muchos de ellos todavía sin identificar. Con el fin de investigar el posible papel de TCERG1 en la biogénesis de snRNPs, quisimos analizar si TCERG1 se asocia a conocidos componentes que intervienen en la biosíntesis de los snRNPs, en concreto a CRM1, CBP80, Int11 y SMN. Para realizar estos experimentos, transfectamos células HEK293T con un plásmido que codifica TCERG1 fusionado a un epítipo T7 o bien con un vector T7 vacío como control negativo. Tras inmunoprecipitar con anticuerpo anti-T7 covalentemente unido a resina de agarosa A, las muestras fueron analizadas por WB con anticuerpos específicos dirigidos contra los factores CRM1, CBP80 e Int11. Los datos obtenidos revelaron la presencia de CRM1, CBP80 e Int11 en los inmunoprecipitados, indicando una asociación entre TCERG1 estos factores de la biosíntesis de los snRNPs (Fig. R27-A,B). SMN es otro factor implicado en la biosíntesis de los snRNPs. La interacción entre TCERG1 y SMN *in vitro* ha sido descrita previamente (Cheng et al. 2007). En este estudio quisimos verificar si esta interacción se daba *in vivo* y si la interacción establecida entre estas proteínas era directa o estaba mediada por RNA. Para ello, transfectamos células HEK293T con TCERG1 fusionado a un epítipo T7 o bien con un vector T7 vacío como control negativo. Estas células se lisaron con buffer de lisis T7 para posteriormente, tratar los extractos proteicos con RNAsa A. Como control positivo, células transfectadas con la misma construcción T7-TCERG1, se sometieron a los mismos tratamientos de temperatura pero sin la adición de RNAsa y a continuación, se realizaron ensayos de inmunoprecipitación utilizando

anticuerpos anti-T7 covalentemente unidos a resina de agarosa A sobre estos extractos. Los extractos se analizaron mediante WB utilizando anticuerpos específicos contra SMN. Los datos obtenidos mostraron una interacción entre TCERG1 y SMN sólo cuando las muestras no se trataban con RNAsa (Fig. R27-C), sugiriendo que estas interacciones están siendo estabilizadas por la unión a RNA en las células.

Estos resultados identifican una asociación de TCERG1 con distintas proteínas implicadas en la biogénesis y homeostasis de los CBs de los snRNPs.

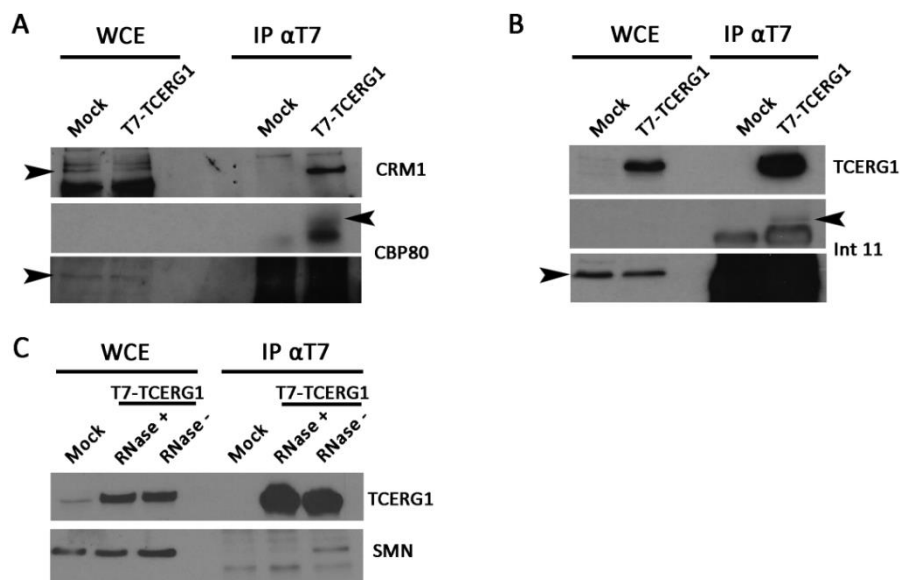


Figura R27. TCERG1 se asocia con distintos componentes de la biosíntesis de los snRNPs. A. Células HEK293T fueron transfectadas transitoriamente con un plásmido que codifica TCERG1 unida al epítipo T7 o un vector vacío como control negativo (Mock). Fracciones del extracto celular total (WCE), fueron analizadas directamente por *western blot* o inmunoprecipitadas con anticuerpos específicos contra T7, seguida de una separación por SDS-PAGE y análisis por WB utilizando anticuerpos específicos para CRM1 y CBP80. Se muestran dos exposiciones diferentes del análisis utilizando el anticuerpo anti-CBP80 (panel inferior). Las flechas indican la posición de las proteínas. **B.** Se realizó el mismo procedimiento experimental descrito en A, pero los extractos fueron analizados con anticuerpos anti-Int11 y anti-TCERG1. Se muestran dos exposiciones diferentes del análisis utilizando el anticuerpo contra Int11 (panel inferior) para su correcta visualización, las flechas indican las bandas correspondientes a la proteína Int11. **C.** Extractos de

células transfectadas con el plásmido pEFBST7-TCERG1 fueron tratados con RNAsa A (0,1 mg/ml, 1h 4°C) o incubadas en las mismas condiciones pero sin la adición de RNAsa para, posteriormente, inmunoprecipitar con anticuerpos específicos contra T7. Se utilizaron anticuerpos específicos contra SMN y TCERG1 para su análisis por WB.

4.2. Efecto de TCERG1 en la formación de snRNPs.

El complejo proceso de la biogénesis de los snRNPs requiere del ensamblaje de siete proteínas Sm alrededor de su sitio de unión que se encuentra en los snRNAs para formar una estructura estable en el citoplasma, seguida por el importe al núcleo y la posterior maduración de los snRNPs (Matera and Wang 2014). Estudios previos del laboratorio han mostrado que TCERG1 interacciona *in vivo* e *in vitro* con las proteínas Sm (Sánchez-Álvarez et al. 2006). Para determinar si TCERG1 está implicado en el proceso de formación de los snRNPs, se llevaron a cabo experimentos de inmunoprecipitación utilizando un anticuerpo que reconoce específicamente a las proteínas Sm (Y12) para cuantificar el RNA asociado a estas proteínas por RT-qPCR. Utilizando las líneas celulares CRISPR HEK293T, observamos que en las líneas 1AC4 y 2AC2, donde TCERG1 se encuentra inhibido, había menor presencia de los snRNAs U1, U2, U4, U5 y U6 en los inmunoprecipitados, indicando una menor unión de esos snRNAs a las proteínas Sm (Fig. R28-A). Resultados similares se obtuvieron al realizar el mismo experimento a partir de extractos celulares procedentes de las líneas HAP1 control y silenciada, con menor presencia de los snRNAs U1, U2, U4, U5 y U6 en los inmunoprecipitados en la línea c012 comparando con la línea control (Fig. R28-B). Para apoyar estos resultados, se realizaron experimentos similares con células HEK293T en condiciones de sobreexpresión de TCERG1. El análisis por RT-qPCR detectó mayor cantidad de snRNAs en los inmunoprecipitados en las muestras donde TCERG1 se encontraba sobreexpresado (Fig. R28-C).

Para descartar la posibilidad de que estos resultados se debieran a que la expresión de TCERG1 estaba afectando la cantidad total de los snRNAs, se cuantificó mediante RT-qPCR la presencia de U1, U2, U4, U5 y U6 tanto en las líneas HEK293T como en HAP1 (Fig. R28-D,E). En ambos casos, no se observó menor cantidad de snRNAs en las líneas silenciadas, de hecho, en el caso de las líneas 1AC4 y 2AC2 se observó un incremento significativo en los niveles de los snRNAs U1 y U5. Y en la línea c012 de las HAP1, se observó mayor cantidad de U1 y U4 respecto a la línea control. Asimismo, al sobreexpresar TCERG1 no se observaron cambios significativos en los niveles de snRNAs entre las células control y las células sobreexpresadas.

Con el fin de comprobar si la inhibición/sobreexpresión de TCERG1 afectaba la expresión de las proteínas Sm realizamos un análisis mediante *western blot* utilizando el anticuerpo Y12, que reconoce las proteínas Sm. Los experimentos indicaron que no se producían cambios en la expresión de las proteínas Sm cuando TCERG1 se encontraba inhibido o sobreexpresado (Fig. R28-G). Además, se comprobó que la cantidad de proteína inmunoprecipitada con el anticuerpo Y12 era similar en las líneas silenciadas respecto a las líneas control (Fig. R28-H). Con estos resultados, concluimos que la expresión de TCERG1 afecta la interacción entre las proteínas Sm y los snRNAs en la formación de las snRNPs.

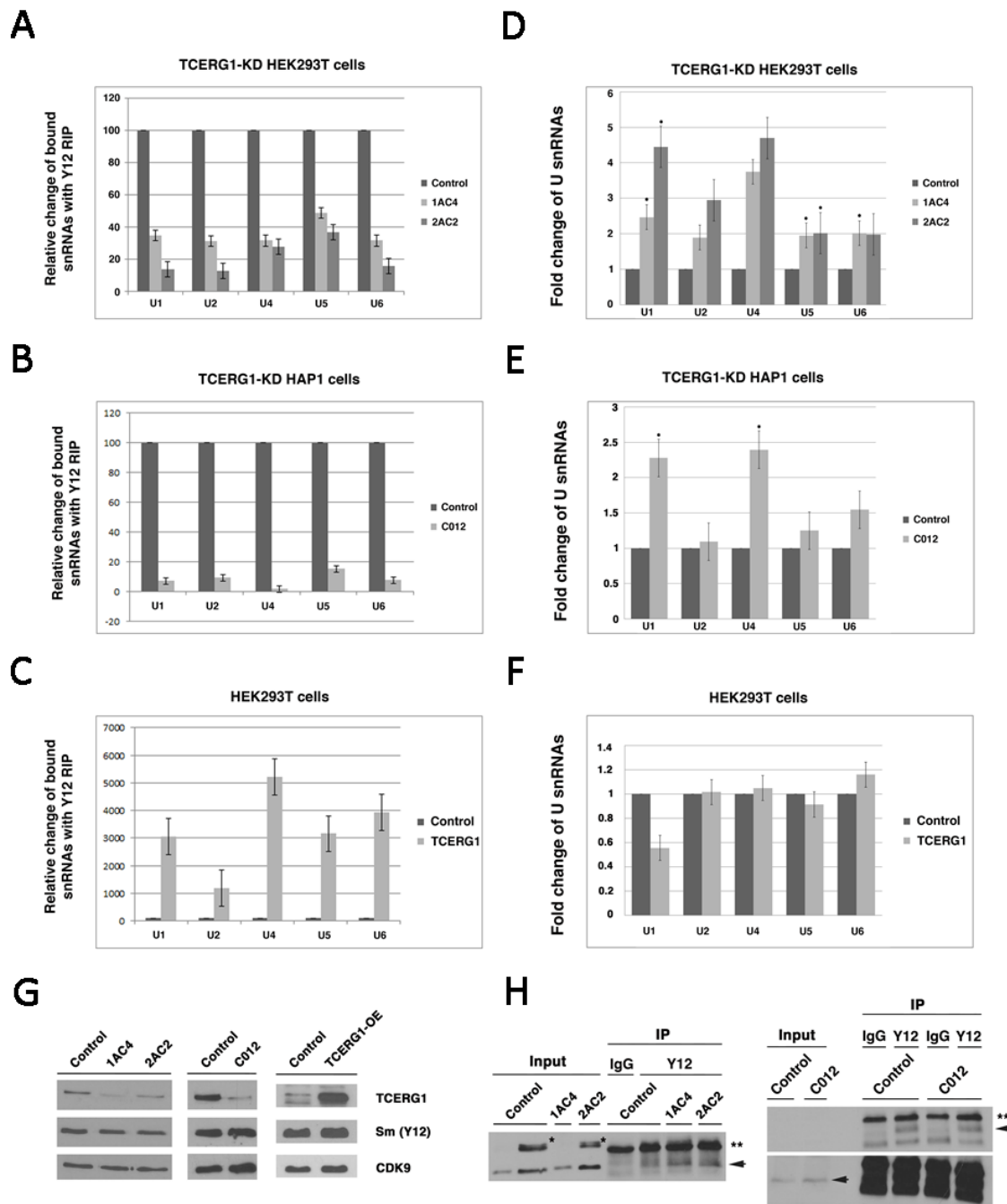


Figura R28. TCERG1 afecta la unión entre las proteínas Sm y los snRNAs. A-B. Extractos celulares de las líneas celulares silenciadas y control de las células HEK293T (A) y HAP1 (B) fueron inmunoprecipitados con el anticuerpo Y12. El RNA unido se extrajo y se cuantificó mediante RT-qPCR. C. El mismo experimento de inmunoprecipitación y análisis por RT-qPCR se llevó a cabo en las líneas HEK293T que habían sido transfectadas con un plásmido de expresión de TCERG1 o el plásmido vacío como control. D-F. Cuantificación del snRNA total de extractos celulares procedentes de las líneas CRISPR HEK293T (D), CRISPR HAP1 (E) y células HEK293T sobreexpresando TCERG1 (F). * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$. G. Análisis de la expresión de las

proteínas Sm mediante *western blot* en las líneas CRISPR HEK293T y CRISPR HAP1 y en las células HEK293T sobreexpresando TCERG1 (OE). El anticuerpo que reconoce la proteína CDK9 se usó como control. **H.** Extractos celulares de las líneas indicadas se inmunoprecipitaron con anti-Sm (Y12) o anti IgG como control negativo y las proteínas se analizaron mediante SDS-PAGE y WB usando el anticuerpo Y12. Una fracción de las muestras se guardó antes de la inmunoprecipitación para su análisis (Input). Se muestran dos exposiciones diferentes del análisis mediante WB de las células HAP1 (derecha) utilizando el anticuerpo anti-Y12 (panel inferior). Las flechas indican las bandas correspondientes a la migración de Sm, el asterisco (*) indica bandas inespecíficas, y el doble asterisco (**) indica las bandas correspondientes a la reacción cruzada de los IgGs en la inmunoprecipitación.

Un defecto en el ensamblaje de los snRNPs debido a la menor expresión de TCERG1 debe reflejarse en una acumulación de los snRNAs U1 y U2 parcialmente ensamblados. Partiendo de esta base, utilizamos anticuerpos que reconocen el extremo 5' trimetilado de los snRNAs (TMG) para realizar experimentos de inmunoprecipitación en las líneas celulares CRISPR HEK293T. Los resultados obtenidos del análisis por RT-qPCR indicaron una mayor cantidad del snRNA U1 precipitado en las líneas 1AC4 y 2AC2, mientras que el nivel del snRNA U2 inmunoprecipitado era similar con respecto a las células control (Fig. R29-A).

Para determinar si TCERG1 influye en la expresión de los precursores de los snRNAs, realizamos nuevos experimentos de RT-qPCR cuantificando los niveles de los precursores de los snRNA U1 y U2 (pre-U1 y pre-U2) en las células con una menor expresión de TCERG1. Los resultados mostraron unos niveles de expresión similares respecto a las células control tanto en las células CRISPR HEK293T como en las células CRISPR HAP1 (Fig. R29-B, C). Estos resultados sugieren que TCERG1 no tiene un efecto directo transcripcional en la expresión de los snRNAs.

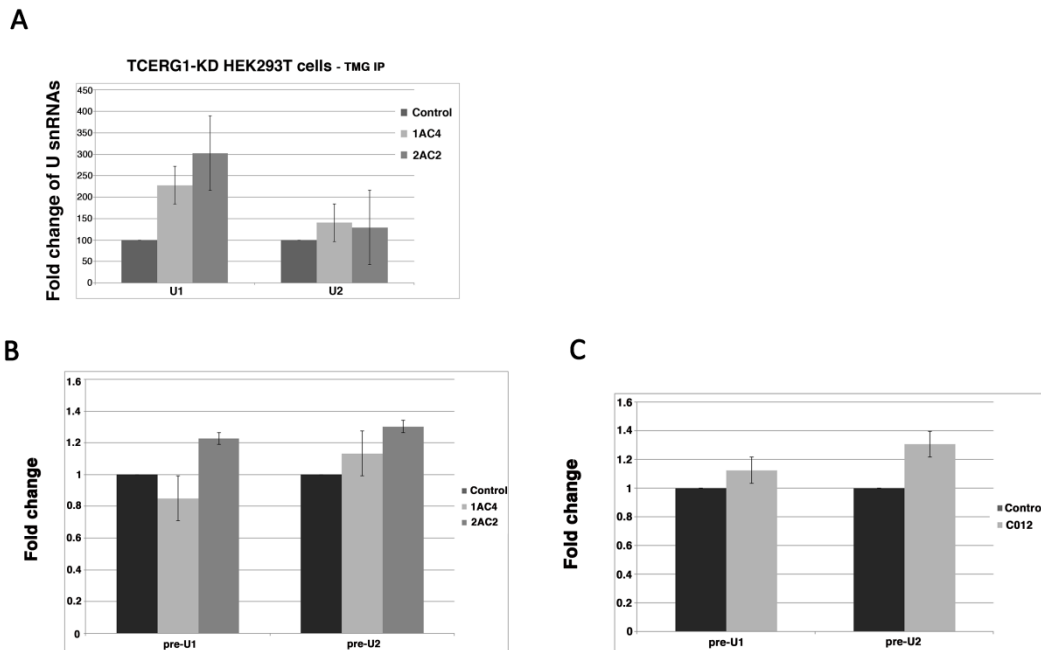


Figura R29. Una menor expresión de TCERG1 da lugar a una acumulación del snRNA U1 parcialmente ensamblado y no tiene efecto en la transcripción de los precursores de los snRNAs U1 y U2. **A.** Extractos celulares de células control y de las líneas 1AC4 y 2AC2 fueron inmunoprecipitados con el anticuerpo que reconoce el TMG. El RNA del precipitado se analizó por RT-qPCR para detectar cambios en los snRNAs U1 y U2. **B, C.** El RNA extraído de las líneas de HEK293T (**A**) y HAP1 (**B**) se analizó mediante RT-qPCR para estudiar los cambios en la expresión de los precursores de los snRNAs U1 y U2 (preU1 y preU2).

4.3. La disminución de la expresión de TCERG1 afecta a la síntesis de snRNPs nacientes.

Los resultados obtenidos anteriormente indican un defecto en el ensamblaje de los snRNPs debido a la menor expresión de TCERG1. Estos resultados pueden deberse a un defecto en la biosíntesis del complejo o bien a una pérdida de estabilidad de los mismos. Para distinguir entre estas posibilidades, analizamos si la producción de snRNPs nacientes se afectaba por una menor expresión de TCERG1. Para ello, contamos con una línea celular U2OS cedida por el doctor Angus I. Lamond (University of Dundee) (Hutten et al. 2014b) que expresa GFP-

SmB bajo el control de doxiciclina. Esta línea celular nos da la ventaja de diferenciar los snRNPs recién sintetizados de los snRNPs presentes ya en la célula.

El análisis mediante *western blot* mostró una expresión de GFP-SmB a las 20h y las 48h de inducción mediante doxiciclina; esta inducción no afectó la expresión de TCERG1 (Fig. R30-A). También comprobamos que GFP-SmB colocalizó correctamente en los CBs y en los *speckles* nucleares mediante IF utilizando anticuerpos que reconocían las proteínas coilina y SRSF2 (rojo) (Fig. R30-B).

Una vez que comprobamos la correcta expresión de GFP-SmB en estas células, realizamos un silenciamiento transitorio con un siRNA control y un siRNA específico para TCERG1 para, a continuación, llevar a cabo una inducción durante 18 horas de doxiciclina. Se prepararon extractos celulares y se realizaron experimentos de inmunoprecipitación utilizando un anticuerpo que reconoce el TMG para precipitar los complejos snRNP/GFP-SmB. Estos extractos se analizaron mediante SDS-PAGE y WB. Mientras GFP-SmB se detectó utilizando un anticuerpo anti-GFP en las células control no pudo ser detectado en los extractos de las células tratadas con el siRNA contra TCERG1. Estos datos sugieren que la producción de nuevos snRNPs es inhibida por la disminución de la expresión de TCERG1.

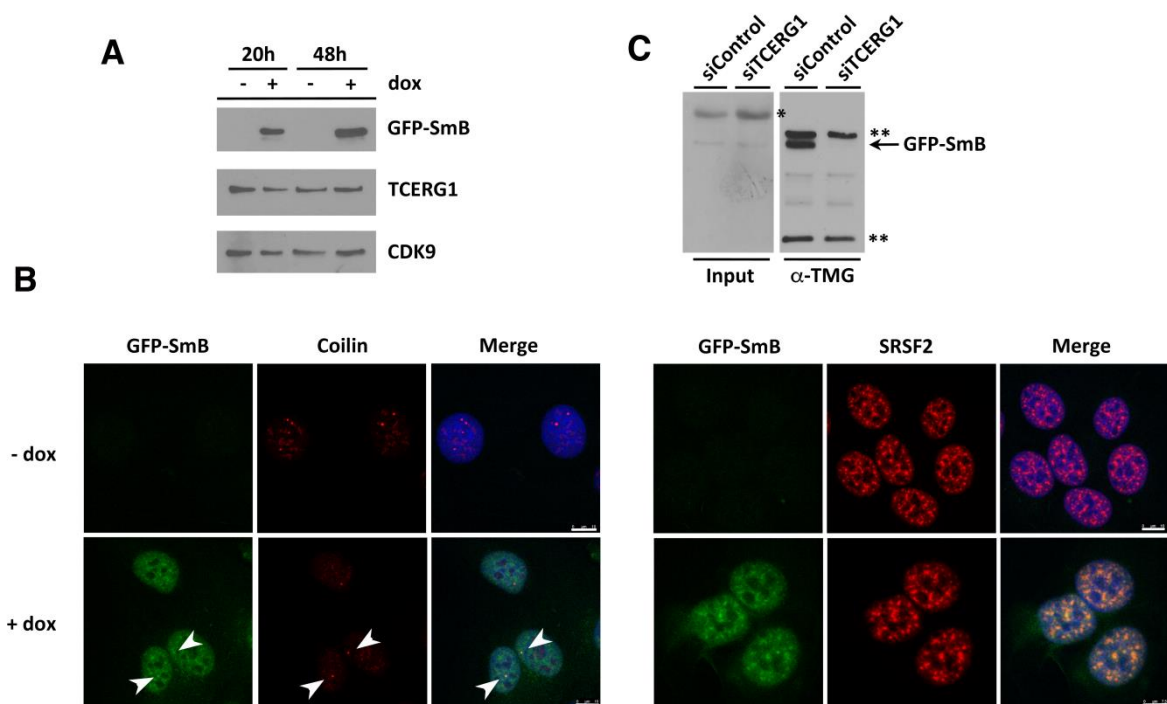


Figura R30. Una menor expresión de TCERG1 afecta a la producción de nuevos snRNPs.

A. Análisis mediante *western blot* de la expresión de GFP-SmB sin (-) y con (+) la adición de doxiciclina durante 20 h y 48 h. Los extractos fueron analizados utilizando anticuerpos específicos contra GFP y TCERG1. El anticuerpo contra CDK9 se usó como control. **B.** La doxiciclina induce la expresión de GFP-SmB (verde), las imágenes muestran los marcajes individuales con coilina (paneles izquierdos) y SRSF2 (paneles derechos) ambas en rojo, así como la superposición de las señales junto con la señal en azul de DAPI (Merge). **C.** Extractos celulares de la línea U2OS transfectada transitoriamente con un siRNA contra TCERG1 o control fueron inducidos con doxiciclina (10 ng/ml) durante 18 h. Después de ese tiempo, una fracción fue analizada por WB (Input) y el resto fue inmunoprecipitada con anticuerpos anti-TMG. A continuación se separaron las proteínas por SDS-PAGE y se analizaron por WB, utilizando anticuerpos anti-GFP.

Conclusiones

1. TCERG1 interacciona con la fosfoproteína NOLC1 *in vivo*. Los dominios FF4 y FF5 de TCERG1 son necesarios para que interaccione con la proteína NOLC1. Dentro de esos dominios, las fenilalaninas 903 y 961 son esenciales para la unión de TCERG1 con NOLC1 *in vivo*. Esta interacción está mediada por RNA.
2. TCERG1 interacciona con coilina y se localiza en la periferia de los CBs en células.
3. La depleción de TCERG1 afecta la integridad de los CBs. Una disminución de la expresión de TCERG1 da lugar a un menor número y tamaño de CBs por núcleo celular. La menor expresión de TCERG1 no afecta a la expresión de coilina. El efecto de TCERG1 sobre los CBs es específico ya que la reexpresión de TCERG1 rescata el fenotipo de las células silenciadas.
4. La sobreexpresión saturante de TCERG1 afecta de forma específica a la integridad de los CBs. Las fenilalaninas 903 y 961 de los dominios FF4 y FF5 median este efecto. La sobreexpresión de TCERG1 no afecta a la expresión de coilina.
5. La inhibición o la sobreexpresión saturante de TCERG1 no afecta significativamente a otros orgánulos nucleares como los *speckles*, *gems*, HLB o el nucléolo.
6. TCERG1 autointeracciona *in vivo* y puede formar homocomplejos.
7. El silenciamiento de NOLC1 o coilina no afecta a la localización de TCERG1, pero sí afecta a la integridad de los CBs. No afecta el fenotipo de otros orgánulos nucleares como *speckles*, *gems*, o nucléolo.

8. TCERG1 interacciona con distintas proteínas implicadas en la biogénesis y homeostasis de los CBs de los snRNPs como CRM1, CBP80 e Int11 y SMN. La interacción de TCERG1 y SMN está mediada por RNA.
9. La inhibición de la expresión de TCERG1 afecta a la unión de las proteínas Sm a los snRNAs y por tanto, afecta a la formación de los complejos snRNPs. En cambio, una sobreexpresión de TCERG1 favorece la unión de las proteínas Sm a los snRNAs.
10. Ni una disminución de la expresión, ni una sobreexpresión de TCERG1 afecta a la cantidad total de snRNAs.
11. TCERG1 no tiene un efecto directo transcripcional significativo en la expresión de los snRNAs, al medir los niveles de pre-U1 y pre-U2 en las líneas celulares silenciadas de TCERG1.
12. La disminución de la expresión de TCERG1 afecta a la síntesis de snRNPs nacientes.

Discusión

Los principales hallazgos de esta tesis doctoral son dos: TCERG1 tiene un papel importante en 1) la integridad de los CBs, y en 2) el ensamblaje del complejo formado por las proteínas Sm y los snRNAs durante el proceso de biogénesis de los snRNPs.

Esta tesis se inició buscando factores asociados a la señal de localización a los *speckles* nucleares de TCERG1. Estudios previos habían demostrado que los dominios FF4 y FF5 de TCERG1 eran los responsables de la localización de TCERG1 a la periferia de los *speckles* nucleares. Dentro de esos dominios, los residuos de fenilalanina F-903 y F-961 eran esenciales para la integridad de los mismos y, por ende, de la señal de localización nuclear (Sánchez-Hernández et al. 2012). Posteriormente se identificó una asociación específica de estos dominios con la proteína nucleolar NOLC1 *in vitro*. Esta asociación parecía ser específica para la versión silvestre de los dominios FF4/FF5, ya que mutaciones en los residuos F-903 y F-961 de esos dominios impiden dicha asociación ((Sánchez-Hernández 2013), ver Introducción).

Los experimentos iniciales realizados en esta tesis han corroborado la asociación específica entre los dominios FF4 y FF5, que comprenden la secuencia desde el aminoácido 878 hasta el aminoácido 1021, y NOLC1 en células (Figura R1). Dado que NOLC1 se localiza, además de en el nucléolo, en los CBs, no fue una sorpresa observar que TCERG1 se localiza en la periferia de esos cuerpos nucleares (Figura R7). Además TCERG1 interacciona con coilina (Fig. R6 - C) y ésta a su vez con NOLC1 (Fig. R6 - B), por lo que estas proteínas podrían estar formando un heterocomplejo.

1. Diferencias en la expresión de TCERG1 alteran la integridad de los CBs

Los resultados obtenidos nos hicieron investigar en profundidad la relación de TCERG1 con los CBs. Los CBs son orgánulos nucleares conservados, involucrados en las fases nucleares de la biogénesis de los snRNPs, elementos esenciales para que se dé lugar el *splicing*. A pesar de haber sido descubiertos hace más de 100 años, los estudios sobre la formación y la estabilidad de los CBs comenzaron hace unas dos décadas y todavía no ha sido dilucidado en detalle hasta la fecha. En esta tesis mostramos que TCERG1 afecta la integridad o formación de los CBs. La inhibición de TCERG1 en varias líneas celulares produce la disgregación de los CBs, medida por la deslocalización de su componente estructural coilina (Figuras R10, R12, R14). Este efecto es específico ya que la sobreexpresión controlada de TCERG1 restaura la arquitectura de los CBs en las líneas celulares silenciadas (Figura R17). Además, el efecto de la inhibición de TCERG1 sobre los CBs parece ser único para estos cuerpos nucleares, ya que no afecta a la estructura de otros orgánulos nucleares, quizás con la excepción de los *speckles* (Figuras R15, R16 – C, D). El efecto de la inhibición de TCERG1 sobre los *speckles* no es total en el sentido de que desaparezcan, pero sí que se aprecia un cambio en su morfología ya que parecen ser de menor tamaño y más difusamente distribuidos (Figuras R9- D, R11). Será interesante continuar investigando el papel de TCERG1 en la formación de los *speckles* en el futuro.

El efecto de la inhibición de TCERG1 es reminiscente del efecto que se observa cuando no hay coilina en distintos organismos (Collier et al. 2006; Liu et al. 2009; Strzelecka et al. 2010; Tucker et al. 2001). La inhibición de otras proteínas involucradas en la maduración de los snRNPs, como SMN, TGS1, PHAX (Girard et al. 2006; Lemm et al. 2006), INTS4 (Takata et al. 2012), WRAP53 (Mahmoudi et al. 2010) o USPL1 (Hutten et al. 2014a) provoca el mismo efecto sobre los CBs.

Estos datos sugieren la presencia de complejos multiproteicos que regulan la integridad o formación de los CBs. De hecho, hemos podido comprobar en esta tesis como TCERG1 se asocia a varios de los componentes que regulan la biogénesis de los snRNPs, como el factor del complejo *Integrator* INT11, la proteína de unión al extremo *cap* CBP80 y la proteína CRM1 (Figura R27), implicada en el transporte del núcleo al citoplasma de diferentes clases de RNAs celulares. La inhibición de NOLC1 (al igual que TCERG1) deslocaliza a la coilina (Figura R25), apoyando la existencia de un complejo multiproteico importante para la homeostasis de los CBs. La presencia de NOLC1 en los CBs puede estar relacionada con la maduración de los snoRNPs que también tiene lugar en los CBs, por lo que se concentran distintos factores implicados en su biosíntesis y procesamiento (Boulon et al. 2004). Es interesante señalar que la pérdida de la localización de NOLC1 en los CBs se ha relacionado con la severidad de la atrofia muscular espinal (SMA) (Renvoise et al. 2009).

Por último, no podemos olvidar el posible papel del RNA en el establecimiento de estas interacciones. El tratamiento con RNAsa A disminuyó la asociación de TCERG1 con NOLC1 (Figura R5), lo que sugiere que esta interacción es dependiente de la integridad del RNA, bien de forma directa o mediada por otras proteínas unidas al RNA. El tratamiento con RNAsa A no modificó el efecto de la inhibición de TCERG1 en el fenotipo de los CBs (Figura R24), sugiriendo que una vez formado el orgánulo el RNA ya no es necesario para el mantenimiento de su integridad.

¿Cuál es el mecanismo por el cual TCERG1 afecta a la integridad de los CBs? Hemos comprobado que TCERG1 está presente en la periferia de los CBs, pero no en su núcleo o *core*, (Figura R7), sugiriendo que ni es un componente constitutivo de los CBs ni asiste a coilina como una proteína de anclaje para la formación de los CBs. Una posibilidad es que TCERG1 pueda estar controlando la expresión de algún componente fundamental de este cuerpo nuclear, dado su papel en la

transcripción y el *splicing* de numerosos genes celulares (Muñoz-Cobo et al. 2017; Pearson et al. 2008). También nos fijamos en la interacción de TCERG1 con SMN, proteína que ha sido implicada en el ensamblaje del anillo Sm y cuya inhibición produce defectos en la formación de los CBs (Raimer, Gray, and Matera 2017). Sin embargo, nuestros resultados muestran que la localización de SMN en *Gems* no está afectada por diferencias en la expresión de TCERG1 (Fig R15 –A, B), por lo que es poco probable que SMN sea el factor involucrado en la desestabilización de los CBs con la inhibición de TCERG1.

Una tercera posibilidad se centra en el papel de los snRNPs como elementos estructurales de los CBs. La formación de los CBs depende de la presencia de los snRNPs (Lemm et al. 2006; Sleeman, Ajuh, and Lamond 2001). Además, snRNPs inmaduros inducen la formación de CBs en células donde no existían estas estructuras, mientras snRNPs maduros no generan nuevos CBs (Novotny et al. 2015; Roithova et al. 2018). Estos datos apoyan un modelo de autoensamblaje a través de la coilina, en el que diferentes especies de RNA nuclea CBs inmaduros, llamados sub-CBs, que se fusionan para ensamblar CBs maduros (Machyna, Heyn, and Neugebauer 2013). TCERG1 podría alterar la formación o arquitectura de estos orgánulos interfiriendo con los sub-CBs dependientes de snRNPs por inhibir la formación de los complejos Sm-snRNP (ver más adelante). De hecho, hemos observado que existe una mayor colocalización de TCERG1 y coilina en células en fase G1, donde la transcripción de los genes de los snRNAs es más elevada, por lo que hay un mayor reclutamiento de los mismos hacia la periferia de los CBs, incrementando la tasa de biogénesis de los snRNPs.

Paradójicamente, la sobreexpresión saturante de TCERG1 tiene los mismos efectos en el fenotipo de los CBs que la inhibición de la expresión de TCERG1 (Figura R18). Al igual que sucede con la inhibición de la expresión de TCERG1, la sobreexpresión de este factor no afecta el fenotipo de otros orgánulos nucleares (Figura R21). El fenómeno de la disgregación específica de los CBs en condiciones

tanto de inhibición como de sobreexpresión saturante ha sido también descrita con la coilina y con WRAP53, un componente de los CBs (Hebert and Matera 2000; Mahmoudi et al. 2010). Dados estos resultados, planteamos la hipótesis de que el fenotipo producido por la sobreexpresión de TCERG1 (o de los factores anteriormente mencionados) pudiera ser debido a un efecto de interferencia negativa entre proteínas exógenas (es decir, ectópicamente expresadas) y endógenas. Si esto fuera así, TCERG1 debería de interactuar consigo mismo en las células. Los resultados obtenidos en los experimentos de sobreexpresión de distintas construcciones de TCERG1 con distintos epítomos seguidos de inmunoprecipitación con anticuerpos específicos confirmó la interacción de las proteínas sobreexpresadas y con la proteína endógena, apoyando el modelo de “secuestro” de TCERG1 endógeno que conduciría a la pérdida de la integridad de los CBs. Claramente, hacen falta más experimentos para entender el mecanismo de este fenómeno que ocurre con varias proteínas importantes para la homeostasis de los CBs.

2. Papel de TCERG1 en la biosíntesis de los snRNPs

El segundo hallazgo importante y novedoso de esta tesis es la implicación de TCERG1 en la biogénesis de los snRNPs. El defecto en la asociación de las proteínas Sm con los snRNAs para formar el complejo snRNP lo hemos observado combinando estrategias de inhibición de la expresión de TCERG1 con experimentos de inmunoprecipitación (Fig. R28 – A, B). La formación de este complejo tras la inhibición de TCERG1 se ve también disminuida al monitorizar la expresión de una proteína de fusión GFP-SmB (Fig. R30). El papel de TCERG1 en la formación del complejo Sm-snRNP no parece implicar un efecto sobre la transcripción de los snRNAs dado que, en las condiciones experimentales usadas, los niveles de los snRNAs endógenos no están disminuidos (Fig R28 D, E). De

hecho, los niveles de los snRNAs endógenos se encuentran aumentados por la inhibición de TCERG1, lo que es sorprendente dada la inhibición del ensamblaje de los snRNPs. A falta de una explicación basada en datos experimentales, intentamos explicar este hecho como un mecanismo compensatorio como respuesta a las perturbaciones en la formación de complejos snRNP maduros, ya sea porque los snRNAs son más estables o porque sus vías de degradación están siendo afectadas. Las observaciones descritas de que los complejos snRNPs, pero no los snRNAs, sean capaces de generar nuevos CBs después de su depleción y que el anillo de Sm sea suficiente para dirigir a los snRNPs a los CBs (Roithova et al. 2018) estarían de acuerdo con el modelo anteriormente expuesto en el cual la inhibición de TCERG1 interferiría con la formación de complejos sub_CBs.

En el proceso de su biogénesis, los snRNAs se mueven desde el núcleo al citoplasma y regresan a los CBs nucleares antes de dirigirse y almacenarse en los *speckles*, desde donde salen para participar en el proceso de *splicing* del pre-mRNA. TCERG1 se identificó como una proteína de localización nuclear (Suñé et al. 1997) y trabajos posteriores, incluyendo esta tesis, han confirmado esta localización (Sánchez-Álvarez et al. 2006; Sánchez-Hernández et al. 2012). Sin embargo, un trabajo reciente ha identificado la presencia de TCERG1 en el citoplasma de células neuronales (Muñoz-Cobo et al. 2017), por lo que planteamos que TCERG1 pueda desplazarse del núcleo al citoplasma para modular la formación de los complejos Sm-snRNPs. Los datos ya publicados mostrando interacciones de TCERG1 con proteínas Sm (Sánchez-Álvarez et al. 2006) y los mostrados en esta tesis acerca de la asociación de TCERG1 con coilina (Fig R7) y su implicación en la biogénesis de los snRNPs apoyan esta hipótesis.

En el futuro será interesante profundizar en el papel de TCERG1 en la biogénesis de los snRNPs y en la formación de los CBs para relacionarlo con la función de TCERG1 en la regulación del *splicing* alternativo.

Bibliografía

- Adriaens, C., L. Standaert, J. Barra, M. Latil, A. Verfaillie, P. Kalev, B. Boeckx, P. W. Wijnhoven, E. Radaelli, W. Vermi, E. Leucci, G. Lapouge, B. Beck, J. van den Oord, S. Nakagawa, T. Hirose, A. A. Sablina, D. Lambrechts, S. Aerts, C. Blanpain, and J. C. Marine. 2016. 'p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity', *Nat Med*, 22: 861-8.
- Alm-Kristiansen, A. H., T. Saether, V. Matre, S. Gilfillan, O. Dahle, and O. S. Gabrielsen. 2008. 'FLASH acts as a co-activator of the transcription factor c-Myb and localizes to active RNA polymerase II foci', *Oncogene*, 27: 4644-56.
- Alvarez, M., X. Estivill, and S. de la Luna. 2003. 'DYRK1A accumulates in splicing speckles through a novel targeting signal and induces speckle disassembly', *J Cell Sci*, 116: 3099-107.
- Allen, M., A. Friedler, O. Schon, and M. Bycroft. 2002. 'The structure of an FF domain from human HYPA/FBP11', *J Mol Biol*, 323: 411-6.
- Andersen, J. S., C. E. Lyon, A. H. Fox, A. K. Leung, Y. W. Lam, H. Steen, M. Mann, and A. I. Lamond. 2002. 'Directed proteomic analysis of the human nucleolus', *Curr Biol*, 12: 1-11.
- Andrade, L. E., E. K. Chan, I. Raska, C. L. Peebles, G. Roos, and E. M. Tan. 1991. 'Human autoantibody to a novel protein of the nuclear coiled body: immunological characterization and cDNA cloning of p80-coilin', *J Exp Med*, 173: 1407-19.
- Baillat, D., M. A. Hakimi, A. M. Naar, A. Shilatifard, N. Cooch, and R. Shiekhattar. 2005. 'Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II', *Cell*, 123: 265-76.
- Banani, S. F., A. M. Rice, W. B. Peeples, Y. Lin, S. Jain, R. Parker, and M. K. Rosen. 2016. 'Compositional Control of Phase-Separated Cellular Bodies', *Cell*, 166: 651-63.
- Barcaroli, D., D. Dinsdale, M. H. Neale, L. Bongiorno-Borbone, M. Ranalli, E. Munarriz, A. E. Sayan, J. M. McWilliam, T. M. Smith, E. Fava, R. A. Knight, G. Melino, and V. De Laurenzi. 2006. 'FLASH is an essential component of Cajal bodies', *Proc Natl Acad Sci U S A*, 103: 14802-7.
- Bartova, E., V. Foltankova, S. Legartova, P. Sehnalova, D. V. Sorokin, J. Suchankova, and S. Kozubek. 2014. 'Coilin is rapidly recruited to UVA-induced DNA lesions and gamma-radiation affects localized movement of Cajal bodies', *Nucleus*, 5: 460-8.
- Becerra, S., M. Montes, C. Hernández-Munain, and C. Suñé. 2015. 'Prp40 pre-mRNA processing factor 40 homolog B (PRPF40B) associates with SF1 and U2AF(65) and modulates alternative pre-mRNA splicing in vivo', *RNA*, 21: 438-57.
- Bedford, M. T., and P. Leder. 1999. 'The FF domain: a novel motif that often accompanies WW domains', *Trends Biochem Sci*, 24: 264-5.

- Bell, M., S. Schreiner, A. Damianov, R. Reddy, and A. Bindereif. 2002. 'p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor', *EMBO J*, 21: 2724-35.
- Berciano, M. T., M. Novell, N. T. Villagra, I. Casafont, R. Bengoechea, J. F. Val-Bernal, and M. Lafarga. 2007. 'Cajal body number and nucleolar size correlate with the cell body mass in human sensory ganglia neurons', *J Struct Biol*, 158: 410-20.
- Berget, S. M., C. Moore, and P. A. Sharp. 1977. 'Spliced segments at the 5' terminus of adenovirus 2 late mRNA', *Proc Natl Acad Sci U S A*, 74: 3171-5.
- Bernardi, R., and P. P. Pandolfi. 2007. 'Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies', *Nat Rev Mol Cell Biol*, 8: 1006-16.
- Biggiogera, M., K. Burki, S. H. Kaufmann, J. H. Shaper, N. Gas, F. Amalric, and S. Fakan. 1990. 'Nucleolar distribution of proteins B23 and nucleolin in mouse preimplantation embryos as visualized by immunoelectron microscopy', *Development*, 110: 1263-70.
- Bohmann, K., J. A. Ferreira, and A. I. Lamond. 1995. 'Mutational analysis of p80 coilin indicates a functional interaction between coiled bodies and the nucleolus', *J Cell Biol*, 131: 817-31.
- Boisvert, F. M., J. Cote, M. C. Boulanger, P. Cleroux, F. Bachand, C. Autexier, and S. Richard. 2002. 'Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing', *J Cell Biol*, 159: 957-69.
- Boisvert, F. M., S. van Koningsbruggen, J. Navascues, and A. I. Lamond. 2007. 'The multifunctional nucleolus', *Nat Rev Mol Cell Biol*, 8: 574-85.
- Bongiorno-Borbone, L., A. De Cola, P. Vernole, L. Finos, D. Barcaroli, R. A. Knight, G. Melino, and V. De Laurenzi. 2008. 'FLASH and NPAT positive but not Coilin positive Cajal Bodies correlate with cell ploidy', *Cell Cycle*, 7: 2357-67.
- Boon, K. L., R. J. Grainger, P. Ehsani, J. D. Barrass, T. Auchynnikava, C. F. Inglehearn, and J. D. Beggs. 2007. 'prp8 mutations that cause human retinitis pigmentosa lead to a U5 snRNP maturation defect in yeast', *Nat Struct Mol Biol*, 14: 1077-83.
- Boulon, S., C. Verheggen, B. E. Jady, C. Girard, C. Pescia, C. Paul, J. K. Ospina, T. Kiss, A. G. Matera, R. Bordonne, and E. Bertrand. 2004. 'PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli', *Mol Cell*, 16: 777-87.
- Broome, H. J., Z. I. Carrero, H. E. Douglas, and M. D. Hebert. 2013. 'Phosphorylation regulates coilin activity and RNA association', *Biol Open*, 2: 407-15.
- Burbelo, P. D., S. Miyamoto, A. Utani, S. Brill, K. M. Yamada, A. Hall, and Y. Yamada. 1995. 'p190-B, a new member of the Rho GAP family, and Rho are induced to cluster after integrin cross-linking', *J Biol Chem*, 270: 30919-26.

- Burlet, P., C. Huber, S. Bertrand, M. A. Ludosky, I. Zwaenepoel, O. Clermont, J. Roume, A. L. Delezoide, J. Cartaud, A. Munnich, and S. Lefebvre. 1998. 'The distribution of SMN protein complex in human fetal tissues and its alteration in spinal muscular atrophy', *Hum Mol Genet*, 7: 1927-33.
- Caceres, J. F., T. Misteli, G. R. Srean, D. L. Spector, and A. R. Krainer. 1997. 'Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity', *J Cell Biol*, 138: 225-38.
- Cajal, S Ramón. 1903. 'Un sencillo metodo de coloracion selectiva del reticulo protoplasmatico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados', *Trab. Lab. Invest. Biol.(Madrid)*, 2: 129-221.
- Carmo-Fonseca, M., J. Ferreira, and A. I. Lamond. 1993. 'Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis--evidence that the coiled body is a kinetic nuclear structure', *J Cell Biol*, 120: 841-52.
- Carvalho, T., F. Almeida, A. Calapez, M. Lafarga, M. T. Berciano, and M. Carmo-Fonseca. 1999. 'The spinal muscular atrophy disease gene product, SMN: A link between snRNP biogenesis and the Cajal (coiled) body', *J Cell Biol*, 147: 715-28.
- Cioce, M., S. Boulon, A. G. Matera, and A. I. Lamond. 2006. 'UV-induced fragmentation of Cajal bodies', *J Cell Biol*, 175: 401-13.
- Clelland, A. K., N. P. Kinnear, L. Oram, J. Burza, and J. E. Sleeman. 2009. 'The SMN protein is a key regulator of nuclear architecture in differentiating neuroblastoma cells', *Traffic*, 10: 1585-98.
- Collier, S., A. Pendle, K. Boudonck, T. van Rij, L. Dolan, and P. Shaw. 2006. 'A distant coilin homologue is required for the formation of cajal bodies in Arabidopsis', *Mol Biol Cell*, 17: 2942-51.
- Cullen, B. R. 2003. 'Nuclear RNA export', *J Cell Sci*, 116: 587-97.
- Charroux, B., L. Pellizzoni, R. A. Perkinson, A. Shevchenko, M. Mann, and G. Dreyfuss. 1999. 'Gemin3: A novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems', *J Cell Biol*, 147: 1181-94.
- Chen, H. K., C. Y. Pai, J. Y. Huang, and N. H. Yeh. 1999. 'Human Nopp140, which interacts with RNA polymerase I: implications for rRNA gene transcription and nucleolar structural organization', *Mol Cell Biol*, 19: 8536-46.
- Chen, J., N. Ezzeddine, B. Waltenspiel, T. R. Albrecht, W. D. Warren, W. F. Marzluff, and E. J. Wagner. 2012. 'An RNAi screen identifies additional members of the Drosophila Integrator complex and a requirement for cyclin C/Cdk8 in snRNA 3'-end formation', *RNA*, 18: 2148-56.

- Chen, L. L., and G. G. Carmichael. 2009. 'Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA', *Mol Cell*, 35: 467-78.
- Chen, Y., Z. Deng, S. Jiang, Q. Hu, H. Liu, Z. Songyang, W. Ma, S. Chen, and Y. Zhao. 2015. 'Human cells lacking coilin and Cajal bodies are proficient in telomerase assembly, trafficking and telomere maintenance', *Nucleic Acids Res*, 43: 385-95.
- Cheng, D., J. Cote, S. Shaaban, and M. T. Bedford. 2007. 'The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing', *Mol Cell*, 25: 71-83.
- Darzacq, X., B. E. Jady, C. Verheggen, A. M. Kiss, E. Bertrand, and T. Kiss. 2002. 'Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs', *EMBO J*, 21: 2746-56.
- Derenzini, M., L. Montanaro, and D. Trere. 2009. 'What the nucleolus says to a tumour pathologist', *Histopathology*, 54: 753-62.
- Derenzini, M., D. Trere, A. Pession, L. Montanaro, V. Sirri, and R. L. Ochs. 1998. 'Nucleolar function and size in cancer cells', *Am J Pathol*, 152: 1291-7.
- Deryusheva, S., and J. G. Gall. 2009. 'Small Cajal body-specific RNAs of *Drosophila* function in the absence of Cajal bodies', *Mol Biol Cell*, 20: 5250-9.
- Egloff, S., P. Vitali, M. Tellier, R. Raffel, S. Murphy, and T. Kiss. 2017. 'The 7SK snRNP associates with the little elongation complex to promote snRNA gene expression', *EMBO J*, 36: 934-48.
- Eilbracht, J., and M. S. Schmidt-Zachmann. 2001. 'Identification of a sequence element directing a protein to nuclear speckles', *Proc Natl Acad Sci U S A*, 98: 3849-54.
- Enwerem, II, V. Velma, H. J. Broome, M. Kuna, R. A. Begum, and M. D. Hebert. 2014. 'Coilin association with Box C/D scaRNA suggests a direct role for the Cajal body marker protein in scaRNP biogenesis', *Biol Open*, 3: 240-9.
- Falahati, H., B. Pelham-Webb, S. Blythe, and E. Wieschaus. 2016. 'Nucleation by rRNA Dictates the Precision of Nucleolus Assembly', *Curr Biol*, 26: 277-85.
- Fidan, Z., A. Younis, P. Schmieder, and R. Volkmer. 2011. 'Chemical synthesis of the third WW domain of TCERG 1 by native chemical ligation', *J Pept Sci*, 17: 644-9.
- Fischer, U., and R. Luhrmann. 1990. 'An essential signaling role for the m3G cap in the transport of U1 snRNP to the nucleus', *Science*, 249: 786-90.
- Fornerod, M., M. Ohno, M. Yoshida, and I. W. Mattaj. 1997. 'CRM1 is an export receptor for leucine-rich nuclear export signals', *Cell*, 90: 1051-60.

- Fox, A. H., Y. W. Lam, A. K. Leung, C. E. Lyon, J. Andersen, M. Mann, and A. I. Lamond. 2002. 'Paraspeckles: a novel nuclear domain', *Curr Biol*, 12: 13-25.
- Frey, M. R., and A. G. Matera. 1995. 'Coiled bodies contain U7 small nuclear RNA and associate with specific DNA sequences in interphase human cells', *Proc Natl Acad Sci U S A*, 92: 5915-9.
- — —. 2001. 'RNA-mediated interaction of Cajal bodies and U2 snRNA genes', *J Cell Biol*, 154: 499-509.
- Fuke, H., and M. Ohno. 2008. 'Role of poly (A) tail as an identity element for mRNA nuclear export', *Nucleic Acids Res*, 36: 1037-49.
- Gall, J. G. 2000. 'Cajal bodies: the first 100 years', *Annu Rev Cell Dev Biol*, 16: 273-300.
- Gama-Carvalho, M., R. D. Krauss, L. Chiang, J. Valcarcel, M. R. Green, and M. Carmo-Fonseca. 1997. 'Targeting of U2AF65 to sites of active splicing in the nucleus', *J Cell Biol*, 137: 975-87.
- Ganot, P., B. E. Jady, M. L. Bortolin, X. Darzacq, and T. Kiss. 1999. 'Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA', *Mol Cell Biol*, 19: 6906-17.
- Gao, X., Q. Wang, W. Li, B. Yang, H. Song, W. Ju, S. Liu, and J. Cheng. 2011. 'Identification of nucleolar and coiled-body phosphoprotein 1 (NOLC1) minimal promoter regulated by NF-kappaB and CREB', *BMB Rep*, 44: 70-5.
- Gasch, A., S. Wiesner, P. Martin-Malpartida, X. Ramirez-Espain, L. Ruiz, and M. J. Macias. 2006. 'The structure of Prp40 FF1 domain and its interaction with the crn-TPR1 motif of Clf1 gives a new insight into the binding mode of FF domains', *J Biol Chem*, 281: 356-64.
- Genabai, N. K., S. Ahmad, Z. Zhang, X. Jiang, C. A. Gabaldon, and L. Gangwani. 2015. 'Genetic inhibition of JNK3 ameliorates spinal muscular atrophy', *Hum Mol Genet*, 24: 6986-7004.
- Gerbi, S. A., A. V. Borovjagin, and T. S. Lange. 2003. 'The nucleolus: a site of ribonucleoprotein maturation', *Curr Opin Cell Biol*, 15: 318-25.
- Girard, C., H. Neel, E. Bertrand, and R. Bordonne. 2006. 'Depletion of SMN by RNA interference in HeLa cells induces defects in Cajal body formation', *Nucleic Acids Res*, 34: 2925-32.
- Goldstrohm, A. C., T. R. Albrecht, C. Sune, M. T. Bedford, and M. A. Garcia-Blanco. 2001. 'The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1', *Mol Cell Biol*, 21: 7617-28.

- Goldstrohm, A. C., A. L. Greenleaf, and M. A. Garcia-Blanco. 2001. 'Co-transcriptional splicing of pre-messenger RNAs: considerations for the mechanism of alternative splicing', *Gene*, 277: 31-47.
- Grimm, C., A. Chari, J. P. Pelz, J. Kuper, C. Kisker, K. Diederichs, H. Stark, H. Schindelin, and U. Fischer. 2013. 'Structural basis of assembly chaperone- mediated snRNP formation', *Mol Cell*, 49: 692-703.
- Guo, L., B. I. Giasson, A. Glavis-Bloom, M. D. Brewer, J. Shorter, A. D. Gitler, and X. Yang. 2014. 'A cellular system that degrades misfolded proteins and protects against neurodegeneration', *Mol Cell*, 55: 15-30.
- Hall, L. L., K. P. Smith, M. Byron, and J. B. Lawrence. 2006. 'Molecular anatomy of a speckle', *Anat Rec A Discov Mol Cell Evol Biol*, 288: 664-75.
- Hallais, M., F. Pontvianne, P. R. Andersen, M. Clerici, D. Lener, H. Benbahouche Nel, T. Gostan, F. Vandermoere, M. C. Robert, S. Cusack, C. Verheggen, T. H. Jensen, and E. Bertrand. 2013. 'CBC-ARS2 stimulates 3'-end maturation of multiple RNA families and favors cap-proximal processing', *Nat Struct Mol Biol*, 20: 1358-66.
- Harding, S. M., J. A. Boiarsky, and R. A. Greenberg. 2015. 'ATM Dependent Silencing Links Nucleolar Chromatin Reorganization to DNA Damage Recognition', *Cell Rep*, 13: 251-9.
- Hebert, M. D. 2013. 'Signals controlling Cajal body assembly and function', *Int J Biochem Cell Biol*, 45: 1314-7.
- Hebert, M. D., and A. G. Matera. 2000. 'Self-association of coilin reveals a common theme in nuclear body localization', *Mol Biol Cell*, 11: 4159-71.
- Hebert, M. D., P. W. Szymczyk, K. B. Shpargel, and A. G. Matera. 2001. 'Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein', *Genes Dev*, 15: 2720-9.
- Hedley, M. L., H. Amrein, and T. Maniatis. 1995. 'An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine-rich splicing factor', *Proc Natl Acad Sci U S A*, 92: 11524-8.
- Hein, N., K. M. Hannan, A. J. George, E. Sanij, and R. D. Hannan. 2013. 'The nucleolus: an emerging target for cancer therapy', *Trends Mol Med*, 19: 643-54.
- Henry, R. W., V. Mittal, B. Ma, R. Kobayashi, and N. Hernandez. 1998. 'SNAP19 mediates the assembly of a functional core promoter complex (SNAPc) shared by RNA polymerases II and III', *Genes Dev*, 12: 2664-72.
- Hernandez, N., and A. M. Weiner. 1986. 'Formation of the 3' end of U1 snRNA requires compatible snRNA promoter elements', *Cell*, 47: 249-58.

- Herold, N., C. L. Will, E. Wolf, B. Kastner, H. Urlaub, and R. Luhrmann. 2009. 'Conservation of the protein composition and electron microscopy structure of *Drosophila melanogaster* and human spliceosomal complexes', *Mol Cell Biol*, 29: 281-301.
- Heyn, P., H. Salmonowicz, J. Rodenfels, and K. M. Neugebauer. 2017. 'Activation of transcription enforces the formation of distinct nuclear bodies in zebrafish embryos', *RNA Biol*, 14: 752-60.
- Hirose, T., G. Virnicchi, A. Tanigawa, T. Naganuma, R. Li, H. Kimura, T. Yokoi, S. Nakagawa, M. Benard, A. H. Fox, and G. Pierron. 2014. 'NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies', *Mol Biol Cell*, 25: 169-83.
- Holbert, S., I. Denghien, T. Kiechle, A. Rosenblatt, C. Wellington, M. R. Hayden, R. L. Margolis, C. A. Ross, J. Dausset, R. J. Ferrante, and C. Neri. 2001. 'The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis', *Proc Natl Acad Sci U S A*, 98: 1811-6.
- Hu, D., E. R. Smith, A. S. Garruss, N. Mohaghegh, J. M. Varberg, C. Lin, J. Jackson, X. Gao, A. Saraf, L. Florens, M. P. Washburn, J. C. Eissenberg, and A. Shilatifard. 2013. 'The little elongation complex functions at initiation and elongation phases of snRNA gene transcription', *Mol Cell*, 51: 493-505.
- Hung, K. H., and W. E. Stumph. 2011. 'Regulation of snRNA gene expression by the *Drosophila melanogaster* small nuclear RNA activating protein complex (DmSNAPc)', *Crit Rev Biochem Mol Biol*, 46: 11-26.
- Hutten, S., G. Chachami, U. Winter, F. Melchior, and A. I. Lamond. 2014a. 'A role for the Cajal-body-associated SUMO isopeptidase USPL1 in snRNA transcription mediated by RNA polymerase II', *J Cell Sci*, 127: 1065-78.
- — —. 2014b. 'A role for the Cajal-body-associated SUMO isopeptidase USPL1 in snRNA transcription mediated by RNA polymerase II', *J Cell Sci*, 127: 1065-78.
- Hutten, S., A. Prescott, J. James, S. Riesenberger, S. Boulon, Y. W. Lam, and A. I. Lamond. 2011. 'An intranucleolar body associated with rDNA', *Chromosoma*, 120: 481-99.
- Hwang, Y. C., T. Y. Lu, D. Y. Huang, Y. S. Kuo, C. F. Kao, N. H. Yeh, H. C. Wu, and C. T. Lin. 2009. 'NOLC1, an enhancer of nasopharyngeal carcinoma progression, is essential for TP53 to regulate MDM2 expression', *Am J Pathol*, 175: 342-54.
- Isaac, C., Y. Yang, and U. Thomas Meier. 1998. 'Nopp140 Functions as a Molecular Link Between the Nucleolus and the Coiled Bodies', *J Cell Biol*, 142: 319-29.
- Izaurrealde, E., J. Lewis, C. McGuigan, M. Jankowska, E. Darzynkiewicz, and I. W. Mattaj. 1994. 'A nuclear cap binding protein complex involved in pre-mRNA splicing', *Cell*, 78: 657-68.

- Jablonski, J. A., A. L. Amelio, M. Giacca, and M. Caputi. 2010. 'The transcriptional transactivator Tat selectively regulates viral splicing', *Nucleic Acids Res*, 38: 1249-60.
- Jady, B. E., E. Bertrand, and T. Kiss. 2004. 'Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal', *J Cell Biol*, 164: 647-52.
- Jady, B. E., X. Darzacq, K. E. Tucker, A. G. Matera, E. Bertrand, and T. Kiss. 2003. 'Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm', *EMBO J*, 22: 1878-88.
- Jady, B. E., P. Richard, E. Bertrand, and T. Kiss. 2006. 'Cell cycle-dependent recruitment of telomerase RNA and Cajal bodies to human telomeres', *Mol Biol Cell*, 17: 944-54.
- Jae, N., C. Preusser, T. Kruger, I. D. Tkacz, M. Engstler, S. Michaeli, and A. Bindereif. 2011. 'snRNA-specific role of SMN in trypanosome snRNP biogenesis in vivo', *RNA Biol*, 8: 90-100.
- Jagiello, I., A. Van Eynde, V. Vulsteke, M. Beullens, A. Boudrez, S. Keppens, W. Stalmans, and M. Bollen. 2000. 'Nuclear and subnuclear targeting sequences of the protein phosphatase-1 regulator NIPP1', *J Cell Sci*, 113 Pt 21: 3761-8.
- Jurica, M. S., L. J. Licklider, S. R. Gygi, N. Grigorieff, and M. J. Moore. 2002. 'Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis', *RNA*, 8: 426-39.
- Jurica, M. S., and M. J. Moore. 2003. 'Pre-mRNA splicing: awash in a sea of proteins', *Mol Cell*, 12: 5-14.
- Kaiser, T. E., R. V. Intine, and M. Dundr. 2008. 'De novo formation of a subnuclear body', *Science*, 322: 1713-7.
- Kitao, S., A. Segref, J. Kast, M. Wilm, I. W. Mattaj, and M. Ohno. 2008. 'A compartmentalized phosphorylation/dephosphorylation system that regulates U snRNA export from the nucleus', *Mol Cell Biol*, 28: 487-97.
- Klingauf, M., D. Stanek, and K. M. Neugebauer. 2006. 'Enhancement of U4/U6 small nuclear ribonucleoprotein particle association in Cajal bodies predicted by mathematical modeling', *Mol Biol Cell*, 17: 4972-81.
- Ko, T. K., E. Kelly, and J. Pines. 2001. 'CrkRS: a novel conserved Cdc2-related protein kinase that colocalises with SC35 speckles', *J Cell Sci*, 114: 2591-603.
- Konig, H., N. Matter, R. Bader, W. Thiele, and F. Muller. 2007. 'Splicing segregation: the minor spliceosome acts outside the nucleus and controls cell proliferation', *Cell*, 131: 718-29.

- Lafarga, M., I. Casafont, R. Bengoechea, O. Tapia, and M. T. Berciano. 2009. 'Cajal's contribution to the knowledge of the neuronal cell nucleus', *Chromosoma*, 118: 437-43.
- Lallemand-Breitenbach, V., and H. de The. 2010. 'PML nuclear bodies', *Cold Spring Harb Perspect Biol*, 2: a000661.
- Le, T. T., L. T. Pham, M. E. Butchbach, H. L. Zhang, U. R. Monani, D. D. Covert, T. O. Gavrilina, L. Xing, G. J. Bassell, and A. H. Burghes. 2005. 'SMNDelta7, the major product of the centromeric survival motor neuron (SMN2) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN', *Hum Mol Genet*, 14: 845-57.
- Lefebvre, S., L. Burglen, S. Reboullet, O. Clermont, P. Burlet, L. Viollet, B. Benichou, C. Cruaud, P. Millasseau, M. Zeviani, and et al. 1995. 'Identification and characterization of a spinal muscular atrophy-determining gene', *Cell*, 80: 155-65.
- Lemm, I., C. Girard, A. N. Kuhn, N. J. Watkins, M. Schneider, R. Bordonne, and R. Luhrmann. 2006. 'Ongoing U snRNP biogenesis is required for the integrity of Cajal bodies', *Mol Biol Cell*, 17: 3221-31.
- Lerner, M. R., J. A. Boyle, S. M. Mount, S. L. Wolin, and J. A. Steitz. 1980. 'Are snRNPs involved in splicing?', *Nature*, 283: 220-4.
- Li, H., and P. M. Bingham. 1991. 'Arginine/serine-rich domains of the su(wa) and tra RNA processing regulators target proteins to a subnuclear compartment implicated in splicing', *Cell*, 67: 335-42.
- Lin, K. T., R. M. Lu, and W. Y. Tarn. 2004. 'The WW domain-containing proteins interact with the early spliceosome and participate in pre-mRNA splicing in vivo', *Mol Cell Biol*, 24: 9176-85.
- Lin, Y., D. S. Protter, M. K. Rosen, and R. Parker. 2015. 'Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins', *Mol Cell*, 60: 208-19.
- Lindgren, V., M. Ares, Jr., A. M. Weiner, and U. Francke. 1985. 'Human genes for U2 small nuclear RNA map to a major adenovirus 12 modification site on chromosome 17', *Nature*, 314: 115-6.
- Liu, J. L., C. Murphy, M. Buszczak, S. Clatterbuck, R. Goodman, and J. G. Gall. 2006. 'The Drosophila melanogaster Cajal body', *J Cell Biol*, 172: 875-84.
- Liu, J. L., Z. Wu, Z. Nizami, S. Deryusheva, T. K. Rajendra, K. J. Beumer, H. Gao, A. G. Matera, D. Carroll, and J. G. Gall. 2009. 'Coilin is essential for Cajal body organization in Drosophila melanogaster', *Mol Biol Cell*, 20: 1661-70.
- Liu, Q., and G. Dreyfuss. 1996. 'A novel nuclear structure containing the survival of motor neurons protein', *EMBO J*, 15: 3555-65.

- Liu, Q., U. Fischer, F. Wang, and G. Dreyfuss. 1997. 'The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins', *Cell*, 90: 1013-21.
- Lorson, C. L., E. Hahnen, E. J. Androphy, and B. Wirth. 1999. 'A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy', *Proc Natl Acad Sci U S A*, 96: 6307-11.
- Lukowiak, A. A., A. Narayanan, Z. H. Li, R. M. Terns, and M. P. Terns. 2001. 'The snoRNA domain of vertebrate telomerase RNA functions to localize the RNA within the nucleus', *RNA*, 7: 1833-44.
- Macias, M. J., S. Wiesner, and M. Sudol. 2002. 'WW and SH3 domains, two different scaffolds to recognize proline-rich ligands', *FEBS Lett*, 513: 30-7.
- Machyna, M., P. Heyn, and K. M. Neugebauer. 2013. 'Cajal bodies: where form meets function', *Wiley Interdiscip Rev RNA*, 4: 17-34.
- Machyna, M., S. Kehr, K. Straube, D. Kappei, F. Buchholz, F. Butter, J. Ule, J. Hertel, P. F. Stadler, and K. M. Neugebauer. 2014. 'The coilin interactome identifies hundreds of small noncoding RNAs that traffic through Cajal bodies', *Mol Cell*, 56: 389-99.
- Machyna, M., K. M. Neugebauer, and D. Stanek. 2015. 'Coilin: The first 25 years', *RNA Biol*, 12: 590-6.
- Mahmoudi, S., S. Henriksson, I. Weibrecht, S. Smith, O. Soderberg, S. Stromblad, K. G. Wiman, and M. Farnebo. 2010. 'WRAP53 is essential for Cajal body formation and for targeting the survival of motor neuron complex to Cajal bodies', *PLoS Biol*, 8: e1000521.
- Makarova, O. V., E. M. Makarov, H. Urlaub, C. L. Will, M. Gentzel, M. Wilm, and R. Luhrmann. 2004. 'A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing', *EMBO J*, 23: 2381-91.
- Mannen, T., S. Yamashita, K. Tomita, N. Goshima, and T. Hirose. 2016. 'The Sam68 nuclear body is composed of two RNase-sensitive substructures joined by the adaptor HNRNPL', *J Cell Biol*, 214: 45-59.
- Mao, Y. S., B. Zhang, and D. L. Spector. 2011. 'Biogenesis and function of nuclear bodies', *Trends Genet*, 27: 295-306.
- Marzluff, W. F., and R. J. Duronio. 2002. 'Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences', *Curr Opin Cell Biol*, 14: 692-9.
- Massenet, S., L. Pellizzoni, S. Paushkin, I. W. Mattaj, and G. Dreyfuss. 2002. 'The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway', *Mol Cell Biol*, 22: 6533-41.

- Masuyama, K., I. Taniguchi, N. Kataoka, and M. Ohno. 2004. 'RNA length defines RNA export pathway', *Genes Dev*, 18: 2074-85.
- Matera, A. G., and K. B. Shpargel. 2006. 'Pumping RNA: nuclear bodybuilding along the RNP pipeline', *Curr Opin Cell Biol*, 18: 317-24.
- Matera, A. G., and Z. Wang. 2014. 'A day in the life of the spliceosome', *Nat Rev Mol Cell Biol*, 15: 108-21.
- McAndrew, P. E., D. W. Parsons, L. R. Simard, C. Rochette, P. N. Ray, J. R. Mendell, T. W. Prior, and A. H. Burghes. 1997. 'Identification of proximal spinal muscular atrophy carriers and patients by analysis of SMNT and SMNC gene copy number', *Am J Hum Genet*, 60: 1411-22.
- McCloskey, A., I. Taniguchi, K. Shinmyozu, and M. Ohno. 2012. 'hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export', *Science*, 335: 1643-6.
- Medenbach, J., S. Schreiner, S. Liu, R. Luhrmann, and A. Bindereif. 2004. 'Human U4/U6 snRNP recycling factor p110: mutational analysis reveals the function of the tetratricopeptide repeat domain in recycling', *Mol Cell Biol*, 24: 7392-401.
- Meier, U. T. 1996. 'Comparison of the rat nucleolar protein nopp140 with its yeast homolog SRP40. Differential phosphorylation in vertebrates and yeast', *J Biol Chem*, 271: 19376-84.
- Meier, U. T., and G. Blobel. 1994. 'NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria', *J Cell Biol*, 127: 1505-14.
- Meister, G., D. Buhler, R. Pillai, F. Lottspeich, and U. Fischer. 2001. 'A multiprotein complex mediates the ATP-dependent assembly of spliceosomal U snRNPs', *Nat Cell Biol*, 3: 945-9.
- Meister, G., C. Eggert, D. Buhler, H. Brahm, C. Kambach, and U. Fischer. 2001. 'Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln', *Curr Biol*, 11: 1990-4.
- Miau, L. H., C. J. Chang, W. H. Tsai, and S. C. Lee. 1997. 'Identification and characterization of a nucleolar phosphoprotein, Nopp140, as a transcription factor', *Mol Cell Biol*, 17: 230-9.
- Mintz, P. J., S. D. Patterson, A. F. Neuwald, C. S. Spahr, and D. L. Spector. 1999. 'Purification and biochemical characterization of interchromatin granule clusters', *EMBO J*, 18: 4308-20.
- Mizutani, K., S. Suetsugu, and T. Takenawa. 2004. 'FBP11 regulates nuclear localization of N-WASP and inhibits N-WASP-dependent microspike formation', *Biochem Biophys Res Commun*, 313: 468-74.

- Montes, M., S. Becerra, M. Sanchez-Alvarez, and C. Sune. 2012. 'Functional coupling of transcription and splicing', *Gene*, 501: 104-17.
- Montes, M., A. Cloutier, N. Sanchez-Hernandez, L. Michelle, B. Lemieux, M. Blanchette, C. Hernandez-Munain, B. Chabot, and C. Sune. 2012. 'TCERG1 regulates alternative splicing of the Bcl-x gene by modulating the rate of RNA polymerase II transcription', *Mol Cell Biol*, 32: 751-62.
- Morris, D. P., and A. L. Greenleaf. 2000. 'The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II', *J Biol Chem*, 275: 39935-43.
- Mouaikel, J., U. Narayanan, C. Verheggen, A. G. Matera, E. Bertrand, J. Tazi, and R. Bordonne. 2003. 'Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron', *EMBO Rep*, 4: 616-22.
- Muñoz-Cobo, Juan Pablo, Noemí Sánchez-Hernández, Sara Gutiérrez, Younes El Yousfi, Marta Montes, Carme Gallego, Cristina Hernández-Munain, and Carlos Suñé. 2017. 'Transcriptional Elongation Regulator 1 Affects Transcription and Splicing of Genes Associated with Cellular Morphology and Cytoskeleton Dynamics and Is Required for Neurite Outgrowth in Neuroblastoma Cells and Primary Neuronal Cultures', *Molecular Neurobiology*, 54: 7808-23.
- Murphy, M. W., B. L. Olson, and P. G. Siliciano. 2004. 'The yeast splicing factor Prp40p contains functional leucine-rich nuclear export signals that are essential for splicing', *Genetics*, 166: 53-65.
- Narayanan, U., T. Achsel, R. Luhrmann, and A. G. Matera. 2004. 'Coupled in vitro import of U snRNPs and SMN, the spinal muscular atrophy protein', *Mol Cell*, 16: 223-34.
- Narayanan, U., J. K. Ospina, M. R. Frey, M. D. Hebert, and A. G. Matera. 2002. 'SMN, the spinal muscular atrophy protein, forms a pre-import snRNP complex with snurportin1 and importin beta', *Hum Mol Genet*, 11: 1785-95.
- Nesic, D., G. Tanackovic, and A. Kramer. 2004. 'A role for Cajal bodies in the final steps of U2 snRNP biogenesis', *J Cell Sci*, 117: 4423-33.
- Neubauer, G., A. King, J. Rappsilber, C. Calvio, M. Watson, P. Ajuh, J. Sleeman, A. Lamond, and M. Mann. 1998. 'Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex', *Nat Genet*, 20: 46-50.
- Novotny, I., M. Blazikova, D. Stanek, P. Herman, and J. Malinsky. 2011. 'In vivo kinetics of U4/U6.U5 tri-snRNP formation in Cajal bodies', *Mol Biol Cell*, 22: 513-23.
- Novotny, I., A. Malinova, E. Stejskalova, D. Mateju, K. Klimesova, A. Roithova, M. Sveda, Z. Knejzlik, and D. Stanek. 2015. 'SART3-Dependent Accumulation of Incomplete Spliceosomal snRNPs in Cajal Bodies', *Cell Rep*, 10: 429-40.

- O'Reilly, D., M. Dienstbier, S. A. Cowley, P. Vazquez, M. Drozd, S. Taylor, W. S. James, and S. Murphy. 2013. 'Differentially expressed, variant U1 snRNAs regulate gene expression in human cells', *Genome Res*, 23: 281-91.
- Ochs, R. L., and R. I. Press. 1992. 'Centromere autoantigens are associated with the nucleolus', *Exp Cell Res*, 200: 339-50.
- Ohno, M., A. Segref, A. Bachi, M. Wilm, and I. W. Mattaj. 2000. 'PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation', *Cell*, 101: 187-98.
- Ohno, M., A. Segref, S. Kuersten, and I. W. Mattaj. 2002. 'Identity elements used in export of mRNAs', *Mol Cell*, 9: 659-71.
- Palacios, I., M. Hetzer, S. A. Adam, and I. W. Mattaj. 1997. 'Nuclear import of U snRNPs requires importin beta', *EMBO J*, 16: 6783-92.
- Palfi, Z., N. Jae, C. Preusser, K. H. Kaminska, J. M. Bujnicki, J. H. Lee, A. Gunzl, C. Kambach, H. Urlaub, and A. Bindereif. 2009. 'SMN-assisted assembly of snRNP-specific Sm cores in trypanosomes', *Genes Dev*, 23: 1650-64.
- Pearson, J. L., T. J. Robinson, M. J. Munoz, A. R. Kornblihtt, and M. A. Garcia-Blanco. 2008. 'Identification of the cellular targets of the transcription factor TCERG1 reveals a prevalent role in mRNA processing', *J Biol Chem*, 283: 7949-61.
- Pelletier, J., G. Thomas, and S. Volarevic. 2018. 'Ribosome biogenesis in cancer: new players and therapeutic avenues', *Nat Rev Cancer*, 18: 51-63.
- Pellizzoni, L., J. Yong, and G. Dreyfuss. 2002. 'Essential role for the SMN complex in the specificity of snRNP assembly', *Science*, 298: 1775-9.
- Pena, E., M. T. Berciano, R. Fernandez, J. L. Ojeda, and M. Lafarga. 2001. 'Neuronal body size correlates with the number of nucleoli and Cajal bodies, and with the organization of the splicing machinery in rat trigeminal ganglion neurons', *J Comp Neurol*, 430: 250-63.
- Pillai, R. S., M. Grimmler, G. Meister, C. L. Will, R. Luhrmann, U. Fischer, and D. Schumperli. 2003. 'Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing', *Genes Dev*, 17: 2321-33.
- Pillai, R. S., C. L. Will, R. Luhrmann, D. Schumperli, and B. Muller. 2001. 'Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein', *EMBO J*, 20: 5470-9.
- Platani, M., I. Goldberg, A. I. Lamond, and J. R. Swedlow. 2002. 'Cajal body dynamics and association with chromatin are ATP-dependent', *Nat Cell Biol*, 4: 502-8.

- Platani, M., I. Goldberg, J. R. Swedlow, and A. I. Lamond. 2000. 'In vivo analysis of Cajal body movement, separation, and joining in live human cells', *J Cell Biol*, 151: 1561-74.
- Polak, P. E., F. Simone, J. J. Kaberlein, R. T. Luo, and M. J. Thirman. 2003. 'ELL and EAF1 are Cajal body components that are disrupted in MLL-ELL leukemia', *Mol Biol Cell*, 14: 1517-28.
- Porokhovnik, L. N., and N. A. Lyapunova. 2019. 'Dosage effects of human ribosomal genes (rDNA) in health and disease', *Chromosome Res*, 27: 5-17.
- Rader, S. D., and C. Guthrie. 2002. 'A conserved Lsm-interaction motif in Prp24 required for efficient U4/U6 di-snRNP formation', *RNA*, 8: 1378-92.
- Raimer, A. C., K. M. Gray, and A. G. Matera. 2017. 'SMN - A chaperone for nuclear RNP social occasions?', *RNA Biol*, 14: 701-11.
- Raska, I., L. E. Andrade, R. L. Ochs, E. K. Chan, C. M. Chang, G. Roos, and E. M. Tan. 1991. 'Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies', *Exp Cell Res*, 195: 27-37.
- Renvoise, B., S. Colasse, P. Burlet, L. Viollet, U. T. Meier, and S. Lefebvre. 2009. 'The loss of the snoRNP chaperone Nopp140 from Cajal bodies of patient fibroblasts correlates with the severity of spinal muscular atrophy', *Hum Mol Genet*, 18: 1181-9.
- Renvoise, B., K. Khoobarly, M. C. Gendron, C. Cibert, L. Viollet, and S. Lefebvre. 2006. 'Distinct domains of the spinal muscular atrophy protein SMN are required for targeting to Cajal bodies in mammalian cells', *J Cell Sci*, 119: 680-92.
- Richard, P., X. Darzacq, E. Bertrand, B. E. Jady, C. Verheggen, and T. Kiss. 2003. 'A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs', *EMBO J*, 22: 4283-93.
- Roithova, A., K. Klimesova, J. Panek, C. L. Will, R. Luhrmann, D. Stanek, and C. Girard. 2018. 'The Sm-core mediates the retention of partially-assembled spliceosomal snRNPs in Cajal bodies until their full maturation', *Nucleic Acids Res*, 46: 3774-90.
- Sahin, U., O. Ferhi, M. Jeanne, S. Benhenda, C. Berthier, F. Jollivet, M. Niwa-Kawakita, O. Faklaris, N. Setterblad, H. de The, and V. Lallemand-Breitenbach. 2014. 'Oxidative stress-induced assembly of PML nuclear bodies controls sumoylation of partner proteins', *J Cell Biol*, 204: 931-45.
- Saitoh, N., C. S. Spahr, S. D. Patterson, P. Bubulya, A. F. Neuwald, and D. L. Spector. 2004. 'Proteomic analysis of interchromatin granule clusters', *Mol Biol Cell*, 15: 3876-90.
- Salichs, E., A. Ledda, L. Mularoni, M. M. Alba, and S. de la Luna. 2009. 'Genome-wide analysis of histidine repeats reveals their role in the localization of human proteins to the nuclear speckles compartment', *PLoS Genet*, 5: e1000397.

- Salzler, H. R., D. C. Tatomer, P. Y. Malek, S. L. McDaniel, A. N. Orlando, W. F. Marzluff, and R. J. Duronio. 2013. 'A sequence in the *Drosophila* H3-H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs', *Dev Cell*, 24: 623-34.
- Sánchez-Álvarez, M., M. Montes, N. Sánchez-Hernández, C. Hernández-Munain, and C. Suñé. 2010. 'Differential Effects of Sumoylation on Transcription and Alternative Splicing by Transcription Elongation Regulator 1 (TCERG1)', *J Biol Chem*, 285: 15220-33.
- Sanchez-Alvarez, M., N. Sanchez-Hernandez, and C. Sune. 2011. 'Spatial Organization and Dynamics of Transcription Elongation and Pre-mRNA Processing in Live Cells', *Genet Res Int*, 2011: 626081.
- Sánchez-Álvarez, Miguel, Aaron C. Goldstrohm, Mariano A. Garcia-Blanco, and Carlos Suñé. 2006. 'Human Transcription Elongation Factor CA150 Localizes to Splicing Factor-Rich Nuclear Speckles and Assembles Transcription and Splicing Components into Complexes through Its Amino and Carboxyl Regions', *Molecular and Cellular Biology*, 26: 4998-5014.
- Sánchez-Hernández, Noemí. 2013. 'Estudio de la señal de localización y comportamiento dinámico del factor de transcripción y splicing TCERG1', Granada.
- Sánchez-Hernández, Noemí, Lidia Ruiz, Miguel Sánchez-Álvarez, Marta Montes, Maria J. Macias, Cristina Hernández-Munain, and Carlos Suñé. 2012. 'The FF4 and FF5 Domains of Transcription Elongation Regulator 1 (TCERG1) Target Proteins to the Periphery of Speckles', *The Journal of Biological Chemistry*, 287: 17789-800.
- Schaffert, N., M. Hossbach, R. Heintzmann, T. Achsel, and R. Luhrmann. 2004. 'RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies', *EMBO J*, 23: 3000-9.
- Scherer, M., and T. Stamminger. 2016. 'Emerging Role of PML Nuclear Bodies in Innate Immune Signaling', *J Virol*, 90: 5850-54.
- Schwartz, J. C., X. Wang, E. R. Podell, and T. R. Cech. 2013. 'RNA seeds higher-order assembly of FUS protein', *Cell Rep*, 5: 918-25.
- Singh, R., and R. Reddy. 1989. 'Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA', *Proc Natl Acad Sci U S A*, 86: 8280-3.
- Sleeman, J. E., P. Ajuh, and A. I. Lamond. 2001. 'snRNP protein expression enhances the formation of Cajal bodies containing p80-coilin and SMN', *J Cell Sci*, 114: 4407-19.
- Sleeman, J. E., L. Trinkle-Mulcahy, A. R. Prescott, S. C. Ogg, and A. I. Lamond. 2003. 'Cajal body proteins SMN and Coilin show differential dynamic behaviour in vivo', *J Cell Sci*, 116: 2039-50.

- Smith, E. R., C. Lin, A. S. Garrett, J. Thornton, N. Mohaghegh, D. Hu, J. Jackson, A. Saraf, S. K. Swanson, C. Seidel, L. Florens, M. P. Washburn, J. C. Eissenberg, and A. Shilatifard. 2011. 'The little elongation complex regulates small nuclear RNA transcription', *Mol Cell*, 44: 954-65.
- Smith, K. P., and J. B. Lawrence. 2000. 'Interactions of U2 gene loci and their nuclear transcripts with Cajal (coiled) bodies: evidence for PreU2 within Cajal bodies', *Mol Biol Cell*, 11: 2987-98.
- Smith, M. J., S. Kulkarni, and T. Pawson. 2004. 'FF domains of CA150 bind transcription and splicing factors through multiple weak interactions', *Mol Cell Biol*, 24: 9274-85.
- Spector, D. L., X. D. Fu, and T. Maniatis. 1991. 'Associations between distinct pre-mRNA splicing components and the cell nucleus', *EMBO J*, 10: 3467-81.
- Spector, D. L., and A. I. Lamond. 2011. 'Nuclear speckles', *Cold Spring Harb Perspect Biol*, 3.
- Stanek, D., and K. M. Neugebauer. 2004. 'Detection of snRNP assembly intermediates in Cajal bodies by fluorescence resonance energy transfer', *J Cell Biol*, 166: 1015-25.
- Stanek, D., J. Pridalova-Hnilicova, I. Novotny, M. Huranova, M. Blazikova, X. Wen, A. K. Saprà, and K. M. Neugebauer. 2008. 'Spliceosomal small nuclear ribonucleoprotein particles repeatedly cycle through Cajal bodies', *Mol Biol Cell*, 19: 2534-43.
- Stanek, D., S. D. Rader, M. Klingauf, and K. M. Neugebauer. 2003. 'Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies', *J Cell Biol*, 160: 505-16.
- Stern, J. L., K. G. Zyner, H. A. Pickett, S. B. Cohen, and T. M. Bryan. 2012. 'Telomerase recruitment requires both TCAB1 and Cajal bodies independently', *Mol Cell Biol*, 32: 2384-95.
- Strzelecka, M., S. Trowitzsch, G. Weber, R. Luhrmann, A. C. Oates, and K. M. Neugebauer. 2010. 'Coilin-dependent snRNP assembly is essential for zebrafish embryogenesis', *Nat Struct Mol Biol*, 17: 403-9.
- Suñé, C., T. Hayashi, Y. Liu, W. S. Lane, R. A. Young, and M. A. Garcia-Blanco. 1997. 'CA150, a nuclear protein associated with the RNA polymerase II holoenzyme, is involved in Tat-activated human immunodeficiency virus type 1 transcription', *Mol Cell Biol*, 17: 6029-39.
- Suñé, Carlos, and Mariano A. Garcia-Blanco. 1999. 'Transcriptional Cofactor CA150 Regulates RNA Polymerase II Elongation in a TATA-Box-Dependent Manner', *Molecular and Cellular Biology*, 19: 4719-28.
- Suzuki, T., H. Izumi, and M. Ohno. 2010. 'Cajal body surveillance of U snRNA export complex assembly', *J Cell Biol*, 190: 603-12.

- Takata, H., H. Nishijima, K. Maeshima, and K. Shibahara. 2012. 'The integrator complex is required for integrity of Cajal bodies', *J Cell Sci*, 125: 166-75.
- Thiry, M. 1993. 'Differential location of nucleic acids within interchromatin granule clusters', *Eur J Cell Biol*, 62: 259-69.
- Thomas Meier, U., and Günter Blobel. 1992. 'Nopp 140 shuttles on tracks between nucleolus and cytoplasm', *Cell*, 70: 127-38.
- Tiku, V., and A. Antebi. 2018. 'Nucleolar Function in Lifespan Regulation', *Trends Cell Biol*, 28: 662-72.
- Tkacz, I. D., Y. Lustig, M. Z. Stern, M. Biton, M. Salmon-Divon, A. Das, V. Bellofatto, and S. Michaeli. 2007. 'Identification of novel snRNA-specific Sm proteins that bind selectively to U2 and U4 snRNAs in *Trypanosoma brucei*', *RNA*, 13: 30-43.
- Tomlinson, R. L., J. Li, B. R. Culp, R. M. Terns, and M. P. Terns. 2010. 'A Cajal body-independent pathway for telomerase trafficking in mice', *Exp Cell Res*, 316: 2797-809.
- Tomlinson, R. L., T. D. Ziegler, T. Supakorndej, R. M. Terns, and M. P. Terns. 2006. 'Cell cycle-regulated trafficking of human telomerase to telomeres', *Mol Biol Cell*, 17: 955-65.
- Toyota, C. G., M. D. Davis, A. M. Cosman, and M. D. Hebert. 2010. 'Coilin phosphorylation mediates interaction with SMN and SmB', *Chromosoma*, 119: 205-15.
- Trinkle-Mulcahy, L., S. Boulon, Y. W. Lam, R. Urcia, F. M. Boisvert, F. Vandermoere, N. A. Morrice, S. Swift, U. Rothbauer, H. Leonhardt, and A. Lamond. 2008. 'Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes', *J Cell Biol*, 183: 223-39.
- Tsekrekou, M., K. Stratigi, and G. Chatzinikolaou. 2017. 'The Nucleolus: In Genome Maintenance and Repair', *Int J Mol Sci*, 18.
- Tucker, K. E., M. T. Berciano, E. Y. Jacobs, D. F. LePage, K. B. Shpargel, J. J. Rossire, E. K. Chan, M. Lafarga, R. A. Conlon, and A. G. Matera. 2001. 'Residual Cajal bodies in coilin knockout mice fail to recruit Sm snRNPs and SMN, the spinal muscular atrophy gene product', *J Cell Biol*, 154: 293-307.
- Tycowski, K. T., M. D. Shu, A. Kukoyi, and J. A. Steitz. 2009. 'A conserved WD40 protein binds the Cajal body localization signal of scaRNP particles', *Mol Cell*, 34: 47-57.
- Tycowski, K. T., Z. H. You, P. J. Graham, and J. A. Steitz. 1998. 'Modification of U6 spliceosomal RNA is guided by other small RNAs', *Mol Cell*, 2: 629-38.

- Venteicher, A. S., E. B. Abreu, Z. Meng, K. E. McCann, R. M. Terns, T. D. Veenstra, M. P. Terns, and S. E. Artandi. 2009. 'A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis', *Science*, 323: 644-8.
- Wagner, S., S. Chiosea, and J. A. Nickerson. 2003. 'The spatial targeting and nuclear matrix binding domains of SRm160', *Proc Natl Acad Sci U S A*, 100: 3269-74.
- Walker, M. P., L. Tian, and A. G. Matera. 2009. 'Reduced viability, fertility and fecundity in mice lacking the cajal body marker protein, coilin', *PLoS One*, 4: e6171.
- Wang, Q., I. A. Sawyer, M. H. Sung, D. Sturgill, S. P. Shevtsov, G. Pegoraro, O. Hakim, S. Baek, G. L. Hager, and M. Dundr. 2016. 'Cajal bodies are linked to genome conformation', *Nat Commun*, 7: 10966.
- Warner, J. R. 1999. 'The economics of ribosome biosynthesis in yeast', *Trends Biochem Sci*, 24: 437-40.
- White, A. E., B. D. Burch, X. C. Yang, P. Y. Gasdaska, Z. Dominski, W. F. Marzluff, and R. J. Duronio. 2011. 'Drosophila histone locus bodies form by hierarchical recruitment of components', *J Cell Biol*, 193: 677-94.
- White, R. J. 2011. 'Transcription by RNA polymerase III: more complex than we thought', *Nat Rev Genet*, 12: 459-63.
- Will, C. L., and R. Luhrmann. 2011. 'Spliceosome structure and function', *Cold Spring Harb Perspect Biol*, 3.
- Wu, C. H., and J. G. Gall. 1993. 'U7 small nuclear RNA in C snurposomes of the *Xenopus* germinal vesicle', *Proc Natl Acad Sci U S A*, 90: 6257-9.
- Wu, Z., C. Murphy, and J. G. Gall. 1994. 'Human p80-coilin is targeted to sphere organelles in the amphibian germinal vesicle', *Mol Biol Cell*, 5: 1119-27.
- Xie, W., T. Ling, Y. Zhou, W. Feng, Q. Zhu, H. G. Stunnenberg, I. Grummt, and W. Tao. 2012. 'The chromatin remodeling complex NuRD establishes the poised state of rRNA genes characterized by bivalent histone modifications and altered nucleosome positions', *Proc Natl Acad Sci U S A*, 109: 8161-6.
- Xu, H., R. S. Pillai, T. N. Azzouz, K. B. Shpargel, C. Kambach, M. D. Hebert, D. Schumperli, and A. G. Matera. 2005. 'The C-terminal domain of coilin interacts with Sm proteins and U snRNPs', *Chromosoma*, 114: 155-66.
- Yang, Q., J. Cheng, Y. Liu, Y. Hong, J. J. Wang, and S. L. Zhang. 2004. 'Cloning and identification of NS5ATP2 gene and its spliced variant transactivated by hepatitis C virus non-structural protein 5A', *World J Gastroenterol*, 10: 1735-9.
- Yang, Y., C. Isaac, C. Wang, F. Dragon, V. Pogacic, and U. T. Meier. 2000. 'Conserved composition of mammalian box H/ACA and box C/D small nucleolar

ribonucleoprotein particles and their interaction with the common factor Nopp140', *Mol Biol Cell*, 11: 567-77.

- Yoshida, K., M. Sanada, Y. Shiraishi, D. Nowak, Y. Nagata, R. Yamamoto, Y. Sato, A. Sato-Otsubo, A. Kon, M. Nagasaki, G. Chalkidis, Y. Suzuki, M. Shiosaka, R. Kawahata, T. Yamaguchi, M. Otsu, N. Obara, M. Sakata-Yanagimoto, K. Ishiyama, H. Mori, F. Nolte, W. K. Hofmann, S. Miyawaki, S. Sugano, C. Haferlach, H. P. Koeffler, L. Y. Shih, T. Haferlach, S. Chiba, H. Nakauchi, S. Miyano, and S. Ogawa. 2011. 'Frequent pathway mutations of splicing machinery in myelodysplasia', *Nature*, 478: 64-9.
- Zaric, B., M. Chami, H. Remigy, A. Engel, K. Ballmer-Hofer, F. K. Winkler, and C. Kambach. 2005. 'Reconstitution of two recombinant LSm protein complexes reveals aspects of their architecture, assembly, and function', *J Biol Chem*, 280: 16066-75.
- Zhang, R., B. R. So, P. Li, J. Yong, T. Glisovic, L. Wan, and G. Dreyfuss. 2011. 'Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly', *Cell*, 146: 384-95.
- Zhong, F. L., L. F. Batista, A. Freund, M. F. Pech, A. S. Venteicher, and S. E. Artandi. 2012. 'TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends', *Cell*, 150: 481-94.
- Zhong, F., S. A. Savage, M. Shkreli, N. Giri, L. Jessop, T. Myers, R. Chen, B. P. Alter, and S. E. Artandi. 2011. 'Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita', *Genes Dev*, 25: 11-6.
- Zhu, Y., R. L. Tomlinson, A. A. Lukowiak, R. M. Terns, and M. P. Terns. 2004. 'Telomerase RNA accumulates in Cajal bodies in human cancer cells', *Mol Biol Cell*, 15: 81-90.