



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

ZIENTZIA
ETA TEKNOLOGIA
FAKULTATEA
FACULTAD
DE CIENCIA
Y TECNOLOGÍA



Trabajo Fin de Grado
Grado en Biotecnología

Análisis proteómico de la oncoproteína nucleofosmina

Autora:

Andrea Gutiérrez Fraga

Directora:

Sonia Bañuelos Rodríguez

© 2017, Andrea Gutiérrez Fraga

Leioa, 26 de junio de 2017

ÍNDICE GENERAL

1. INTRODUCCIÓN	1
1.1. LEUCEMIA MIELOIDE AGUDA	1
1.1.1. Descripción y estado actual de la enfermedad	1
1.1.2. Relación entre AML y NPM1	1
1.2. NUCLEOFOSMINA	3
1.2.1. Definición y descripción estructural y funcional de la proteína	3
1.2.2. Mutación de la proteína NPM1. Diferencias entre mutada y salvaje	5
2. HIPÓTESIS Y OBJETIVOS	6
2.1. HIPÓTESIS	6
2.2. OBJETIVOS	7
3. MATERIALES Y MÉTODOS	8
3.1. ELECCIÓN Y CULTIVO DE LAS LÍNEAS CELULARES	8
3.2. ANÁLISIS PROTEÓMICO	8
3.2.1. Fundamento	8
3.2.2. Preparación de las muestras	9
3.2.3. Inmunoprecipitación	9
3.2.4. Análisis estadísticos (espectrometría masas)	10
3.3. EXPERIMENTO F2H	13
3.3.1. Fundamento	13
3.3.2. Preparación de las muestras	14
4. RESULTADOS Y DISCUSIÓN	15
4.1. ANÁLISIS PROTEÓMICO	15
4.2. EXPERIMENTO F2H	20
5. CONCLUSIONES	21
6. BIBLIOGRAFÍA	22
ANEXO I	26

1. INTRODUCCIÓN

1.1. LEUCEMIA MIELOIDE AGUDA

1.1.1. Descripción y estado actual de la enfermedad

La leucemia mieloide aguda (también conocida como mieloblástica aguda, leucemia mielógena aguda, leucemia granulocítica aguda o AML) es un tipo de cáncer que afecta a las células sanguíneas. Se origina en la médula ósea, pero en la mayor parte de casos, se extiende rápidamente a través de la sangre llegando a afectar otras partes del cuerpo como el hígado o el sistema nervioso (American Cancer Society, 2017).

Las personas afectadas por la enfermedad a menudo presentan trastornos/síntomas derivados de una producción insuficiente de células sanguíneas sanas. Los signos más comunes son anemia (bajo recuento de glóbulos rojos), trombocitopenia o déficit de plaquetas, dolor de huesos y articulaciones y sangrado y problemas de coagulación (De Kouchkovsky y Abdul-Hay, 2016).

La leucemia mieloide aguda es una patología poco frecuente en comparación con otros tipos de cáncer, con 3 nuevos casos por cada 100.000 habitantes por año. Sin embargo, la incidencia aumenta progresivamente con la edad, afectando a 13 de cada 100.000 personas de 65 años de edad o mayores. La AML es el tipo de leucemia más común en adultos y representa el 40 % de todas las leucemias en el mundo occidental (Löwenber *et al.*, 1999).

1.1.2. Relación entre AML y NPM1

Si bien se desconoce la causa exacta de la AML, se han detectado alteraciones genéticas en algunos pacientes. El desarrollo de nuevas técnicas de secuenciación masiva en los últimos años, ha permitido la identificación de nuevas mutaciones genéticas asociadas a AML.

Algunos de los genes identificados son: NPM1, KIT, FLT3, CEBPA, BAALC, IDH, RAS, WT1, ERG, MN1, DNMT, TET2, ASXL1, PTPN11 y CBL. De todos ellos destacan NPM1, FLT3 y CEBPA por su asociación con la respuesta al tratamiento y progreso de esta enfermedad (Lagunas- Rangel, 2016).

Concretamente, las mutaciones de la proteína nucleofosmina son las anomalías moleculares más frecuentes en la leucemia mieloide aguda. El gen *NPM1* aparece mutado en aproximadamente el 30 % de los casos de AML (Federici *et al*, 2010).

En 1999, se observaron los primeros casos de AML con NPM1 citoplasmática que no se correspondía a ninguna proteína de fusión de NPM1 anteriormente descrita (Falini *et al*, 2007). La localización típica de la nucleofosmina es el nucléolo, por lo que llamaron a esta variante mutada de la proteína que se encontraba deslocalizada en el citoplasma NPMc+. Esta variante se asocia a un cariotipo normal y es mutuamente excluyente con otras anomalías genéticas asociadas a AML. Esto sugiere que la NPMc+ puede ser el desencadenante de la leucemogénesis, aunque todavía falta por esclarecer completamente el mecanismo de acción.

Las mutaciones en dicho gen ocurren típicamente en el exón 12, afectando concretamente a la región C-terminal de la proteína. Tal mutación del gen se ha descrito como el evento iniciador en un 35 % de casos de leucemia mieloide aguda en adultos. Como resultado, se crea una nueva señal de exportación nuclear (NES). Así, la proteína mutada abandona su localización natural, el nucléolo, y se desplaza al citoplasma (Konoplev *et al*, 2009). Además, el dominio C-terminal de la proteína resulta desestabilizado y parcial o totalmente desplegado, debido a la alteración de su núcleo hidrofóbico (Federici *et al*, 2010).

Estas alteraciones en la proteína mutada provocan una localización citoplasmática aberrante y estable de la proteína en el citosol (Federici *et al*, 2010), un rasgo que podría ser crítico en las transformaciones neoplásicas. En todas las proteínas de fusión de NPM1 asociadas a cáncer, la región N-terminal de la nucleofosmina está conservada, pero supone un dominio fundamental para la acción oncogénica de NPM1 ya que es el sitio de oligomerización (la NPM1 es un pentámero) y de unión para muchos factores y proteínas.

De lo que no hay duda, es que la deslocalización citoplasmática de la proteína es un evento crítico en la patogénesis de la enfermedad. Si bien es cierto que la NPM1 está en continuo movimiento entre nucléolo y citoplasma, la deslocalización citoplasmática de la NPM1 altera diversas vías celulares. Un aumento de la nucleofosmina exportada al citoplasma, puede suponer una pérdida de función (Loss Of Function, LOF), lo que

se refiere a que la NPM1 deja de unir ciertos ligandos nucleolares. También se puede dar una ganancia de función (Gain Of Function, GOF) si el mutante adquiere la capacidad de unir nuevos ligandos en el citoplasma. Además de todo ello, otra posibilidad es que la NPM1 mutante siga uniendo ligandos como la variante salvaje, pero provocando la deslocalización de los mismos, influyendo en su normal funcionamiento (Falini *et al*, 2011).

La valoración del estado de mutación de NPM1 ha mejorado la evaluación del riesgo y elección del tratamiento de la AML. Además de para diagnóstico y evaluación del riesgo, el análisis de la mutación de NPM1 se puede utilizar para supervisar la respuesta a la terapia.

1.2. NUCLEOFOSMINA

1.2.1. Definición y descripción estructural y funcional de la proteína

La nucleofosmina, también conocida como fosfoproteína B23, numatrina o NPM1, es una proteína codificada por el gen *NPM1*. Se trata de una fosfoproteína que pertenece a la familia de chaperonas de la nucleoplasmina. Es de localización fundamentalmente nucleolar, aunque se desplaza constantemente del núcleo al citoplasma y viceversa (Federici *et al*, 2013).

El gen de la NPM1 humana se localiza en el cromosoma 5q35 y está formado por 12 exones, que codifican para tres isoformas (NPM1.1, NPM1.2 y NPM1.3). La transcripción completa del gen *NPM1* resulta en una proteína de 294 aminoácidos con un peso de 32,7 kDa. La nucleofosmina se presenta en forma pentamérica (163,5 kDa) (**Figura 1**) (Box *et al*, 2016).

La nucleofosmina presenta diferentes dominios (**Figura 1**): un dominio N-terminal o *core* bien conservado que adopta una estructura de barril β (Lee *et al*, 2007), un dominio globular C-terminal compuesto por tres hélices α (Grummitt *et al*, 2008) y un dominio central sin estructura regular que actúa a modo de *linker*. La NPM1 alberga en sus diferentes regiones señales de direccionalización que modulan su transporte: motivos de localización nuclear (NLS) y exportación al citoplasma (NES) (**Figura 2**). Su anclaje al nucléolo depende probablemente de la unión a ácidos nucleicos a través del dominio C-terminal.

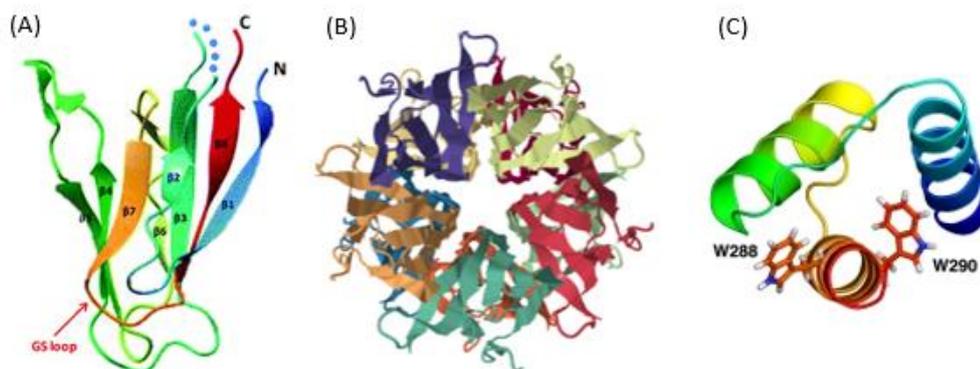


Figura 1. Estructuras tridimensionales de los dominios de la NPM1. (A) Estructura del monómero del *core* (PDB ID 2P1B, Lee *et al*, 2007). (B) Estructura pentamérica del *core* (PDB 2P1B). (C) Estructura del dominio C-terminal. Se resaltan los residuos W288 y W290 esenciales para el plegamiento de la NPM1 (Grummitt *et al*, 2008).

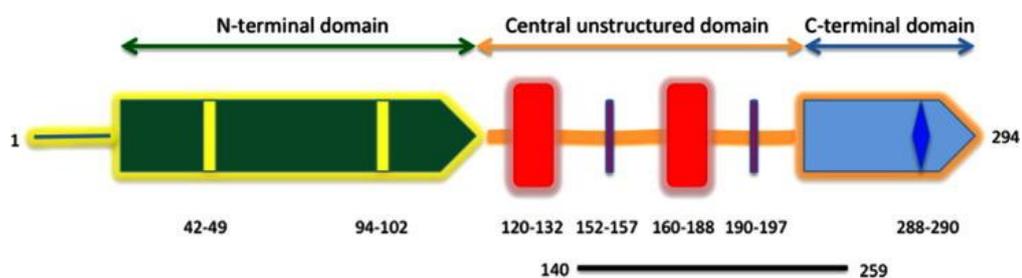


Figura 2. Representación esquemática de los dominios de la NPM1. De izquierda a derecha se representan los dominios N-terminal o *core* (verde) con dos posibles señales de exportación nuclear (NES), reconocidas por el sistema de transporte mediado por la exportina Crm1; la región central no estructurada con la señal de localización nuclear (NLS) bipartita (152-157,190-197), y, por último, el dominio C-terminal (azul) donde se encuentran dos residuos de triptófano (W288 y W290), claves para su plegamiento y la localización nucleolar. La barra negra representa la región con actividad ribonucleasa. (Federici *et al*, 2013).

La región N-terminal está altamente conservada y constituye el dominio central o *core*, el cual es responsable de la oligomerización y las interacciones de NPM1 con otras proteínas como p53, p14arf y APE-1, entre otras. El dominio C-terminal es responsable de la unión de la proteína a ácidos nucleicos (Federici y Falini, 2013).

En cuanto a la función de la nucleofosmina, el mejor descrito es su papel como chaperona de ácidos nucleicos y proteínas. Destaca también su participación en diversos procesos celulares como la biogénesis de ribosomas, duplicación del centrosoma y ensamblaje de histonas (Federici y Falini, 2013).

La NPM1 también cumple un papel crítico en la regulación de la apoptosis. Varios estudios han demostrado la implicación de los niveles de la proteína en el control de la respuesta apoptótica celular (Box *et al*, 2016). La regulación negativa de NPM1 induce la apoptosis, mientras que niveles altos de la proteína protegen contra la apoptosis. Desde el punto de vista de su papel patogénico, niveles elevados de NPM1 podrían promover la transformación tumoral al permitir la supervivencia celular.

Otra de las funciones importantes de la NPM1 es la respuesta al daño del DNA. Recientemente, se ha advertido que la nucleofosmina probablemente juegue un papel directo en la reparación de las lesiones en el DNA, estando implicada en múltiples vías de reparación (Scott y Oeffinger, 2016).

Las rupturas de doble hebra del DNA (Double Strand Breaks) son una de las lesiones más destructivas para las células. Pueden ser reparadas por dos vías: recombinación homóloga (HR) y unión de extremos no homólogos (NHEJ). Se ha visto que la NPM1 está implicada al menos en la recombinación homóloga. Después de la inducción de DSBs, la nucleofosmina sale del nucléolo al nucleoplasma y es reclutada a *foci*, lugares de la cromatina donde se lleva a cabo la reparación. Sin embargo, aún no queda claro el mecanismo exacto de su función en esta vía (Koike *et al*, 2010).

Además de su papel en la reparación de rupturas de doble hebra, la NPM1 también responde a daños del DNA inducidos por luz ultravioleta (UV). Los niveles de nucleofosmina aumentan después de la exposición celular a los rayos UV. Igualmente, la NPM1 también participa en la vía de reparación de escisión de base (BER), inducida tras daño oxidativo. Varias proteínas nucleolares implicadas en BER son deslocalizadas en células deficientes en NPM1 (Box *et al*, 2016).

1.2.2. Mutación de la proteína NPM1. Diferencias entre mutada y salvaje

Las mutaciones en el gen de la nucleofosmina humana son las alteraciones genéticas más frecuentes en la leucemia mieloide aguda. Estas mutaciones resultan en una localización aberrante de la proteína en el citoplasma de las células mieloides.

Se han descrito cerca de 50 variantes mutantes de NPM1 en AML, siendo el mutante A el más común en casos de AML en adultos (aproximadamente un 80%). Éste se forma como consecuencia de la inserción en el gen de un fragmento de 4 nucleótidos,

que provoca un cambio en el marco de lectura, lo que supone la pérdida de residuos esenciales en el plegamiento del dominio C-terminal: los residuos de triptófano W288 y W290 (Falini *et al*, 2011).

Esta mutación es responsable del desplegamiento y la localización aberrante de NPM1 en casos de leucemia mieloide aguda. Se genera un nuevo motivo de exportación nuclear (NES) que favorece la interacción de la nucleofosmina con la exportina Crm1 (Arregi *et al*, 2015). Como consecuencia, se produce una acumulación de NPM1 mutada en el citoplasma de las células leucémicas. Por tanto, la localización aberrante de NPM1 se debe a la acción conjunta de dos factores: el desplegamiento del dominio C-terminal y la adquisición de una nueva NES que promueve la exportación nuclear.

Esta deslocalización es, con mucha probabilidad, la responsable de la alteración de varias vías celulares; ya sea por pérdida de funciones en el nucléolo o por ganancia y aparición de nuevas funciones en el citoplasma. En concreto, podría verse afectada su respuesta al daño del DNA.

Es por ello que un análisis proteómico comparativo de las interacciones de la nucleofosmina salvaje y mutante puede ayudar a esclarecer los mecanismos que desencadenan la enfermedad, permitiendo explorar las posibilidades anteriormente mencionadas (pérdida o ganancia de función y deslocalización de ligandos); siendo especialmente interesantes aquellas interacciones que impliquen proteínas o factores relacionados con reparación de daño del DNA.

2. HIPÓTESIS Y OBJETIVOS

2.1. HIPÓTESIS

La deslocalización de la proteína mutante NPMc+ supone una alteración en las funciones naturales de la célula, debido bien a una pérdida o ganancia de función de la proteína en el nucléolo y citoplasma respectivamente. La interacción alterada de NPMc+ con factores y proteínas implicadas en la regulación del ciclo celular y en vías de reparación de daño afectaría a la proliferación de las células que se vuelve anómala, siendo el principal desencadenante de la AML. Por ello, la identificación de los ligandos implicados en diversos procesos de reparación del DNA que podrían ver afectada su función puede ayudar a discernir los mecanismos moleculares causantes

de tal afección; pudiendo ser un punto de partida para el desarrollo de futuras estrategias terapéuticas para combatir la enfermedad.

2.2. OBJETIVOS

El principal objetivo del proyecto es explorar el papel de la NPM1 en la respuesta al daño al DNA y cómo podría estar alterada dicha función en la AML. Para ello, se llevará a cabo un análisis proteómico comparativo de la NPM1 mutada y salvaje a fin de profundizar en el conocimiento de esta proteína, tanto en condiciones fisiológicas como patológicas. Esto ayudará al diseño de un futuro tratamiento para la enfermedad.

Con el análisis proteómico se pretende identificar ligandos deslocalizados y nuevos de la proteína para explorar las distintas posibilidades de pérdida y ganancia de función, así como deslocalización de ligandos que nos pueden ayudar a entender los mecanismos moleculares que subyacen a la AML. Como sugieren estudios previos del grupo (datos no publicados), la APE1 podría ser uno de los ligandos deslocalizados en AML por lo que hemos orientado parte del trabajo a explorar la interacción entre NPM1 y APE1.

Por tanto, se espera identificar interacciones de la NPM1 con otras proteínas y factores (interactoma) que pueden resultar claves en el desarrollo de AML para poder desarrollar terapias futuras, en las que utilizando algunos de estos ligandos como dianas y biomarcadores se pueda seguir la evolución de la enfermedad.

El trabajo se articulará desarrollando los siguientes objetivos concretos:

- Caracterizar diferencias en el interactoma de la NPM1 salvaje y mutante, que puedan revelar procesos celulares alterados en AML.
- Identificar nuevos ligandos de la NPM1, para profundizar en la funcionalidad de la proteína; en particular, en su papel de reparación.
- Seleccionar ligandos interesantes de la NPM1, para validar y estudiar la interacción que establecen con NPM1.
- Caracterizar la interacción de la NPM1 con uno de sus ligandos: APE1.

3. MATERIALES Y MÉTODOS

3.1. ELECCIÓN Y CULTIVO DE LAS LÍNEAS CELULARES

Para realizar los experimentos que se describen a continuación se trabaja con dos líneas celulares diferentes: células HEK293T para el análisis proteómico de ligandos de la NPM1 y células BHK modificadas (ChromoTek) para el experimento F2H.

Las líneas celulares se cultivan en medio Eagle modificado por Dulbecco (DMEM) y suplementado con suero bovino fetal (FBS) al 10 %, 100 unidades/mL de penicilina y 100 µg/mL de estreptomina (todos de Invitrogen).

Las células se ponen a crecer en frascos (*flasks*) T-25 (BHK) y T-7 (HEK293T) con el medio de cultivo en una estufa a 37 °C y 5% de CO₂. Para hacer los pases, se observan las células al microscopio, y en función de su confluencia se decide la dilución que se va a llevar a cabo. Todos los pasos del protocolo de cultivo se llevan a cabo en condiciones de esterilidad. En primer lugar, se retira el medio viejo para después hacer un lavado con PBS. Tras retirar el PBS se despegan las células con tripsina. Se hace la dilución con la cantidad adecuada de medio fresco y se vuelven a incubar en las condiciones mencionadas hasta el siguiente subcultivo.

3.2. ANÁLISIS PROTEÓMICO

3.2.1. Fundamento

La proteómica nos permite llevar a cabo un estudio a gran escala de todas las proteínas de una célula. En este trabajo se comparan las interacciones de la NPM1 salvaje y mutada. Se identifican los ligandos de la proteína en distintas condiciones para obtener una visión global integrada de los procesos celulares relacionados con la AML.

Para el análisis se aíslan, de las células HEK293T, mediante el método *GFP-Trap*, complejos de nucleofosmina, etiquetada con proteína amarilla fluorescente (YFP-NPM1), tanto salvaje como mutante, además de la etiqueta (solo YFP) como control.

Después se realiza una electroforesis desnaturizante en gel de poliacrilamida al 10% (SDS-PAGE) de las muestras. Tras teñir el gel con Coomassie, se analizan fragmentos del mismo (2 por cada calle) en el Servicio de Proteómica de la Universidad del País Vasco (SGIker) mediante la técnica analítica de espectrometría de masas.

3.2.2. Preparación de las muestras

Para una mayor fiabilidad estadística, se preparan 3 réplicas de las 3 muestras que corresponden a la proteína salvaje fusionada a la etiqueta YFP por el extremo N-terminal (W1, W2 y W3), la proteína mutante con la misma etiqueta (A1, A2 y A3) y la etiqueta sola (Y1, Y2 e Y3). Como ya se ha mencionado, se trabaja con la línea celular HEK293T en este experimento.

En primer lugar, se cuentan las células para sembrarlas en placas Petri. Para el recuento se utilizan una cámara de Neubauer y una máquina de recuento de células automática (BioRad). En el recuento manual se obtuvo una media de 1.24×10^6 células/mL. La medida automática fue de $1,67 \times 10^6$ células/mL. Se tomó como medida de referencia la media del contaje manual. A continuación, se realiza una siembra de $6,5 \times 10^6$ células/mL diluyéndolas con medio fresco en 3 placas Petri.

Al día siguiente, se procede a la transfección. Una de las placas se transfecta con la proteína nucleofosmina salvaje (W), otra con el mutante (A), ambas con la etiqueta YFP, y otra con la etiqueta sola (Y). Se utilizó una preparación de DNA a $0,3 \mu\text{g}/\mu\text{L}$. Para la transfección se añade a cada placa $18 \mu\text{L}$ DNA y $450 \mu\text{L}$ mezcla ($450 \mu\text{L}$ DMEM sin FBS + $22,5 \mu\text{L}$ Xtreme Gene9). Se incuban las placas 25 minutos a temperatura ambiente.

3.2.3. Inmunoprecipitación

Para el aislamiento de los complejos proteicos de NPM1 se realiza una inmunoprecipitación con GFP-Trap. El método consiste en el uso de anticuerpos anti-GFP, que reconocen también la etiqueta YFP, unidos covalentemente a bolas de agarosa magnetizadas (Chromotek) para aislar, mediante un imán, la nucleofosmina etiquetada con la proteína fluorescente y los ligandos que haya podido unir.

24 horas después de la transfección, se retira el medio de las placas Petri y se hace un lavado con 10 mL de PBS frío durante un minuto. Después se retira el PBS y se añade 1 mL tampón de lisis (LB). El tampón LB contiene 10 mL de tampón de lisis (Pierce) al que se le añade un cóctel de inhibidores de proteasas (cOmplete, Roche). Se incuban en hielo durante 30 minutos, resuspendiendo cada 5-7 minutos, y se centrifugan las muestras a 14.000 g 10 minutos.

Durante ese tiempo se equilibran las bolas de agarosa. Se resuspenden en el tampón indicado, y, con ayuda de un imán se sedimentan y se retira dicha disolución. Después se resuspenden las bolitas en 600 μL de LB evitando la formación de espuma, y se dejan 10 minutos en hielo sin el imán.

Para la unión de los complejos (*binding*) se reparten las bolas equitativamente en tres tubos eppendorf y se aplica el imán para que sedimenten y se pueda retirar el tampón. A continuación, se añaden 850 μL de los extractos, reservando una parte (100-150 μL) como control (Whole Cell Extract). Se resuspende y se incuba 3 horas a 4 °C con rotación en noria. Transcurrido ese tiempo, se hacen unos lavados con tampón LB, se recuperan las muestras en 150 μL de tampón de carga para SDS-PAGE, se hierven y se congelan a -20 °C hasta su uso.

Después se realiza una electroforesis desnaturizante SDS-PAGE de las muestras en gel de poliacrilamida al 10%. El gel se deja correr 2 cm a 120 V. Finalmente se tiñe con Coomassie durante al menos 1 hora y se destiñe con una solución decolorante que contiene ácido acético para fijar las proteínas al gel. Se corta el gel con ayuda de un bisturí. En nuestro caso se optó por cortar cada calle en 2 fragmentos para reducir la complejidad de las muestras y facilitar el posterior análisis proteómico (**Figura 5**).

3.2.4. Espectrometría de masas y análisis estadístico

Las bandas del gel se someten a digestión trípica en tampón NH_4HCO_3 50 mM y 12,5 ng/ μl de tripsina (Roche Diagnostics) a 37 °C durante 16 horas. Tras la digestión, se realiza una extracción de péptidos, primero con NH_4HCO_3 25 mM y acetonitrilo y después con 0,1 % de ácido trifluoroacético y acetonitrilo. Las muestras se secan por centrifugación a vacío y se resuspendieron en 13 μl de 0,1 % ácido fórmico, se sonicaron durante 5 min y se centrifugaron a 13.000 rpm durante 10 minutos. El sobrenadante se analiza en un espectrómetro de masas Q Exactive (Thermo Scientific) acoplado a un sistema cromatográfico EASY-nLC 1000 (Thermo Scientific).

Los péptidos digeridos se cargan en una precolumna Acclaim PepMap100 (75 μm x 2 cm, Thermo Scientific) conectada a una columna analítica Acclaim PepMap RSLC (50 μm x 15 cm, Thermo Scientific). Los péptidos se eluyen directamente al nanoES Emitter (Thermo Scientific) mediante un gradiente lineal de 3 % a 30 % de acetonitrilo en 0,1 % de ácido fórmico durante 60 min a un flujo de 300 nl/min.

Los espectros MS se adquirieron a una resolución de 70.000 (m/z 200) y los espectros de fragmentación a 17.500 (m/z 200). Para la selección de los péptidos se utiliza una ventana de 2.0 th y para la fragmentación de los péptidos se aplica una energía de colisión normalizada de 28. El tiempo máximo de inyección fue de 120 ms para las adquisiciones MS y MS/MS y se utilizaron valores de AGC de 3E6 para las adquisiciones MS y de 5E5 para las MS/MS. Los espectros se procesan con MaxQuant (versión 1.5.3.17). Se seleccionó 8 ppm y 20 ppm como tolerancia para los precursores y fragmentos respectivamente y se permitió un máximo de 2 omisiones de corte para la tripsina. Se seleccionó la carbamidometilación de cisteína (Cys) como modificación fija y la oxidación de metionina (Met) y la acetilación N-terminal de proteínas como modificaciones variables. Se utilizó una base de datos SwissProt humana (versión 2017_02) a la que se añadieron las secuencias modificadas de NPM1. Se estableció un 1% de FDR (False Discovery Rate) tanto a nivel de péptidos como de proteínas y se calculó la intensidad normalizada de las proteínas (LFQ) mediante el algoritmo MaxLFQ (Cox *et al*, 2008).

El Servicio de Proteómica de la UPV/EHU nos facilitó los datos de la intensidad normalizada de las proteínas identificadas. Para el análisis cuantitativo de los resultados se utiliza el programa Perseus (Tyanova *et al*, 2016). Este software permite el análisis masivo de datos para extraer información biológicamente significativa a partir de un análisis bioinformático de los datos obtenidos con MaxQuant.

La matriz original con los datos brutos del análisis de espectrometría de masas contenía 3.652 proteínas. Para simplificar la lista y facilitar el posterior análisis de los datos se llevan a cabo varios filtrados (**Figura 3**). Con ellos se elimina la información innecesaria, así como las identificaciones incorrectas de la matriz general de datos. En nuestro caso aplicamos filtros para eliminar los *hits* identificados solamente con péptidos modificados, los reversos y los posibles contaminantes. La matriz resultante tenía 3.443 grupos de proteínas.

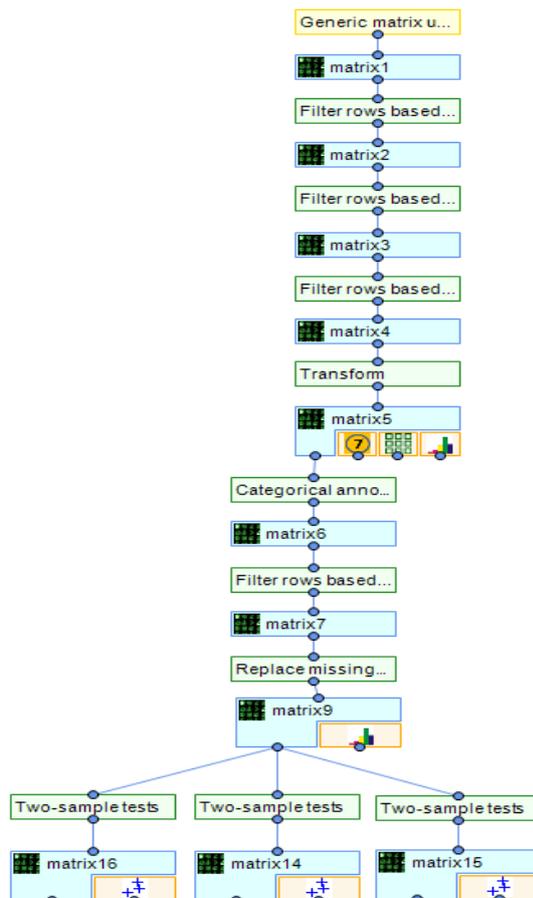


Figura 3. Esquema del proceso de tratamiento de los datos con el programa Perseus. La matriz de datos procedente de MaxQuant (cuadro amarillo) se somete a una serie de filtrados (cuadros verdes) para generar matrices más sencillas (cuadros azules) hasta llegar a las representaciones gráficas (*volcano plots*) de las comparaciones de las muestras.

Para lograr un mejor comportamiento de los datos y poder aplicar los test estadísticos, es adecuada una distribución normal de los datos. Por ello, se transforman a escala logarítmica en base 2, para lograr valores más manejables y que se ajusten a una distribución normal. Esta transformación genera un conjunto de “NaN” (Non Assigned Numbers) que corresponden a valores de expresión iguales a cero, que se generan cuando el software no puede asignar un valor de intensidad de forma fiable. Para comprobar que los datos siguen un comportamiento compatible con una distribución normal se representan los histogramas de cada columna de intensidad correspondiente a cada muestra.

Para definir el rigor del análisis, es necesario fijar el número mínimo válido de valores aceptados en el análisis, esto aumenta la seguridad de los datos. En nuestro caso,

determinamos un número mínimo de 3 valores de intensidad para cada proteína en al menos uno de los grupos (A, W e Y).

Por último, se lleva a cabo la imputación de los valores NaN tomando como referencia los valores por defecto. Perseus analiza la distribución de los datos asumiendo una distribución normal, calcula el ancho y centro de la distribución, y genera valores plausibles para sustituir aquellos valores que faltan. De esta forma se creó una matriz completa que contaba con 2.748 elementos.

A partir de esta última matriz se pueden generar los gráficos de tipo *volcano* que ayudan a una interpretación más visual de los resultados. Con la herramienta *two-sample test* se crean los gráficos comparativos de los parámetros estadísticos *-log p-value* (eje y) y *t-test difference* (eje x), de la nucleofosmina salvaje frente a la etiqueta, W vs Y (**Figura 6**), la nucleofosmina mutante frente a la etiqueta, A vs Y (**Figura 7**) y la proteína salvaje frente al mutante A, W vs A (**Figura 8**). En todos los gráficos se establecieron unos umbrales de significancia para ambos parámetros: 1,33 como valor de referencia para el *-log p-value*, equivalente a un $p < 0,05$ y $\pm 1,0$ para el *t-test difference*, equivalente a un ratio de ± 2 para la expresión de la proteína.

3.3. EXPERIMENTO F2H

Se llevó a cabo un ensayo para detectar la interacción de NPM1 (salvaje y el mutante A) y APE1 en células mediante la técnica F2H (ChromoTek). Las proteínas se cotransfectaron en células BHK (fibroblastos de riñón de crías de hámster) fusionadas a las etiquetas YFP y Cherry. Se utilizó el kit F2H® de la casa comercial ChromoTek.

3.3.1. Fundamento

El F2H es un método desarrollado por la casa comercial ChromoTek para la detección de interacciones proteína-proteína (PPI) en células. La técnica se basa en la expresión de dos proteínas de fusión en una célula eucariota con el genoma modificado de forma que se reclutan las proteínas GFP a uno o varios puntos de la cromatina: una proteína cebo (*bait*) y una proteína presa (*prey*) fusionadas a dos etiquetas fluorescentes (GFP o similar y otra etiqueta roja) (**Figura 4**). La co-localización de la emisión de fluorescencia de ambas proteínas en el núcleo es indicativa de una interacción positiva de las proteínas cebo y presa (Universidad Ludwig-Maximilians, Patente, 2010).

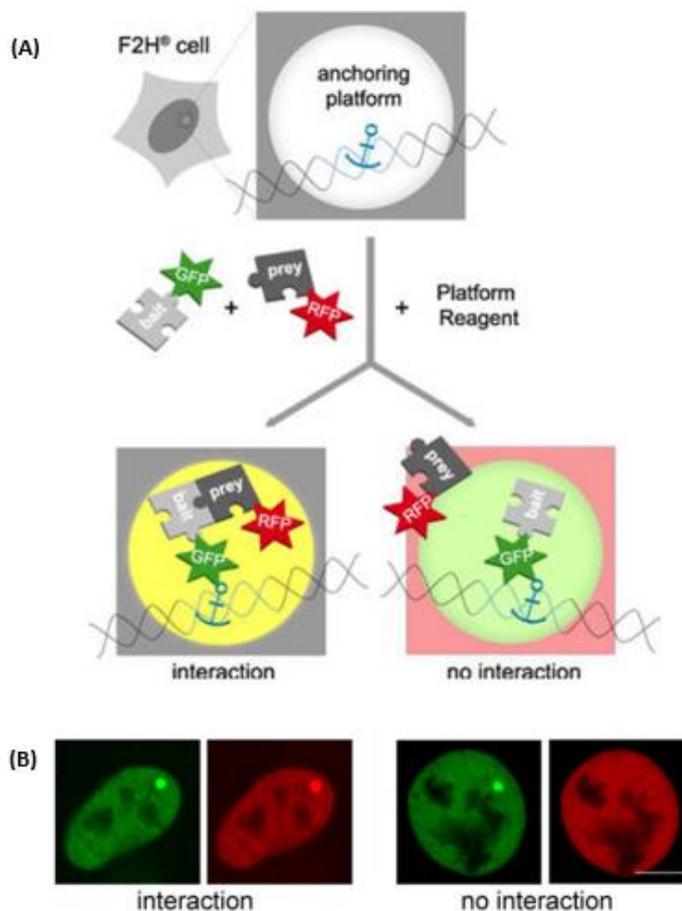


Figura 4. Principio del método F2H. (A) Las células modificadas con la plataforma de anclaje para GFP (*anchoring platform*) se cotransfectan con las proteínas a estudiar (*bait* y *prey*) cada una fusionada con las etiquetas GFP y RFP. (B) La etiqueta GFP se retiene en el punto de anclaje y, si se da la interacción, se obtiene un punto de fluorescencia roja co-localizado con un punto verde de fluorescencia, mientras que si no se da la interacción no aparece dicho punto rojo.

3.3.2. Preparación de las muestras

Se prepararon los plásmidos con las proteínas de fusión: nucleofosmina fusionada a la etiqueta fluorescente amarilla (YFP) por el extremo N-terminal y APE1 fusionada a la etiqueta Cherry por el extremo C-terminal.

Se hace una transfección doble de las células en una placa de pocillos T75 que contiene un cubre siliconizado con las células adheridas. También se transfecta el DNA control del *kit*. Se preparan 3 muestras por duplicado: NPM1 y APE1 con el reactivo “*platform reagent* (PR)” (N/A+), ambas proteínas de fusión sin el reactivo (N/A-) y el control positivo del kit con el reactivo (C+).

La composición de las muestras es la siguiente:

- “N/A – “: 2 μ L NPM + 4 μ L APE1
- “N/A +”: 2 μ L NPM + 4 μ L APE1 + 2 μ L PR
- “C+”: 2 μ L DNA control + 2 μ L PR

Además, se añaden a cada pocillo 80 μ L de la mezcla formada por 240 μ L DMEM y 12 μ L X-tremeGENE 9. Se incuban 25 minutos a temperatura ambiente. Tras 48 horas se fijan las células con paraformaldehído, y se analizan las preparaciones por microscopía de fluorescencia con el microscopio Axio Scope (Zeiss).

4. RESULTADOS Y DISCUSIÓN

4.1. ANÁLISIS PROTEÓMICO

Para la detección de ligandos de la NPM1, se analizan muestras (por triplicado) de YFP-NPM1 WT (“W”), YFP-NPM1mutA (“A”) y YFP (“Y”) sobreexpresadas en HEK293T y purificadas por GFP-Trap.

Las bandas del gel de electroforesis se analizaron por el Servicio de Proteómica de la UPV (SGIker), mediante el método de cuantificación sin etiqueta (*Label Free Quantification*) de espectrometría de masas para determinar la cantidad relativa de proteínas en las distintas muestras biológicas (**Figura 5**).

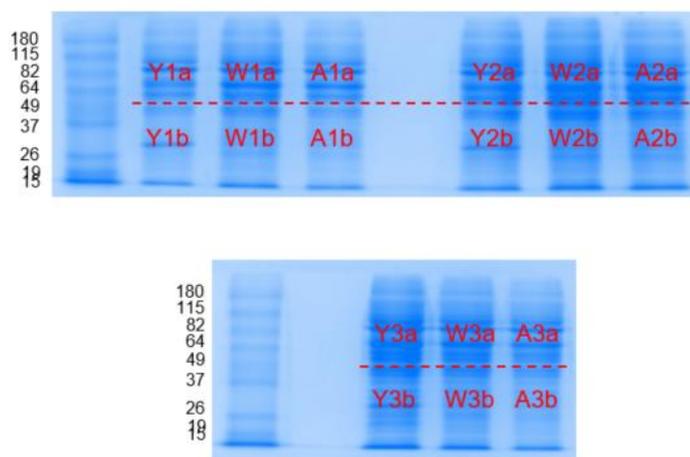


Figura 5. Geles de las muestras GFP-Trap. En la imagen de arriba se muestra el gel con el marcador (1ª calle) y dos de las tres réplicas de las muestras de la proteína salvaje (W), el mutante (A), ambas fusionadas a YFP y la etiqueta sola (Y) como control. En la imagen de abajo se muestra el marcador (1º calle) y la tercera réplica de las muestras. Las líneas de puntos indican cómo se fragmentaron las calles para generar las 18 muestras que se llevaron al Servicio de Proteómica (SGIker).

Los datos obtenidos del análisis de espectrometría de masas con la plataforma MaxQuant, revelaron un listado inicial de 3.652 proteínas, el cual se redujo a 2.748 tras aplicar una serie de filtros de significancia estadística con el programa bioinformático Perseus, como se detalla en el apartado 3.2.4 de materiales y métodos.

Para validar la reproducibilidad o consistencia entre las muestras, se calcula el parámetro estadístico Pearson que refleja la correlación entre las distintas muestras (**Figura 6**). El estadístico muestra una excelente correlación entre las muestras de la proteína salvaje (W1, W2 y W3), la proteína mutada (A1, A2 y A3) y la muestra control de la etiqueta de proteína amarilla fluorescente YFP (Y1, Y2 e Y3). Se obtuvo un valor superior a 0,9 en todos los casos, lo que indica que, en efecto, las muestras son comparables, y las diferencias encontradas entre ellas tienen alta probabilidad de ser significativas.

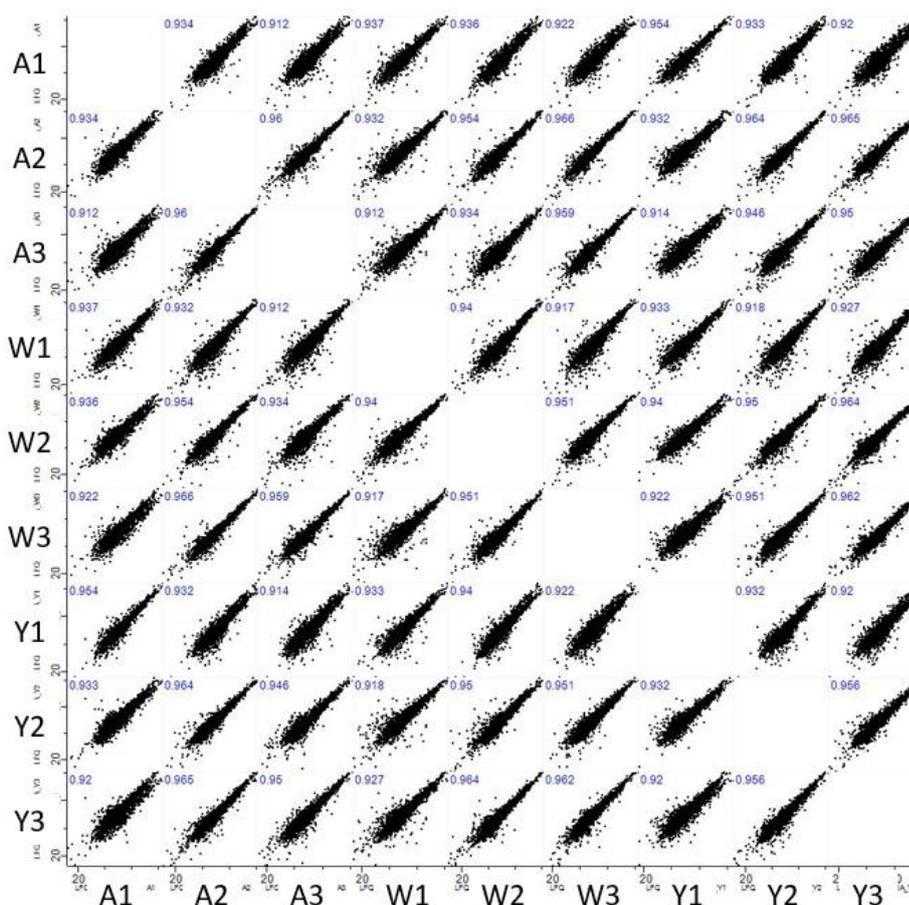


Figura 6. Multi-scatter plot. Gráfico de dispersión múltiple en el que se representan los valores del estadístico de Pearson para la comparación de las 9 muestras entre sí.

Siguiendo el proceso de filtrado explicado anteriormente se originó una matriz final compuesta por 2.748 proteínas. Para un análisis más visual se muestran los resultados en gráficos *volcano plot*. Se generan *volcano plots*, comúnmente utilizados en este tipo de estudios, para las tres posibles comparaciones: la comparación entre la proteína salvaje y el mutante A (**Figura 7**), la etiqueta (Y) frente a la proteína salvaje (W) y el mutante (A) (**Figuras AI.1 y AI.2**).

Para determinar qué proteínas están enriquecidas en las distintas muestras, se establecen umbrales de significancia estadística. Así, según el criterio que hemos utilizado, se consideran proteínas sobrerrepresentadas aquellas que cumplan valores del parámetro $-\log p\text{-value}$ iguales o superiores a 1,3, y, $\pm 1,0$ del parámetro *Student t-test difference*. Por tanto, las proteínas que se encuentran en los cuadrantes superior izquierdo y derecho, dentro de los umbrales, se considera que tienen una expresión diferenciada en las muestras a comparar (**Tablas 1 y 2**).

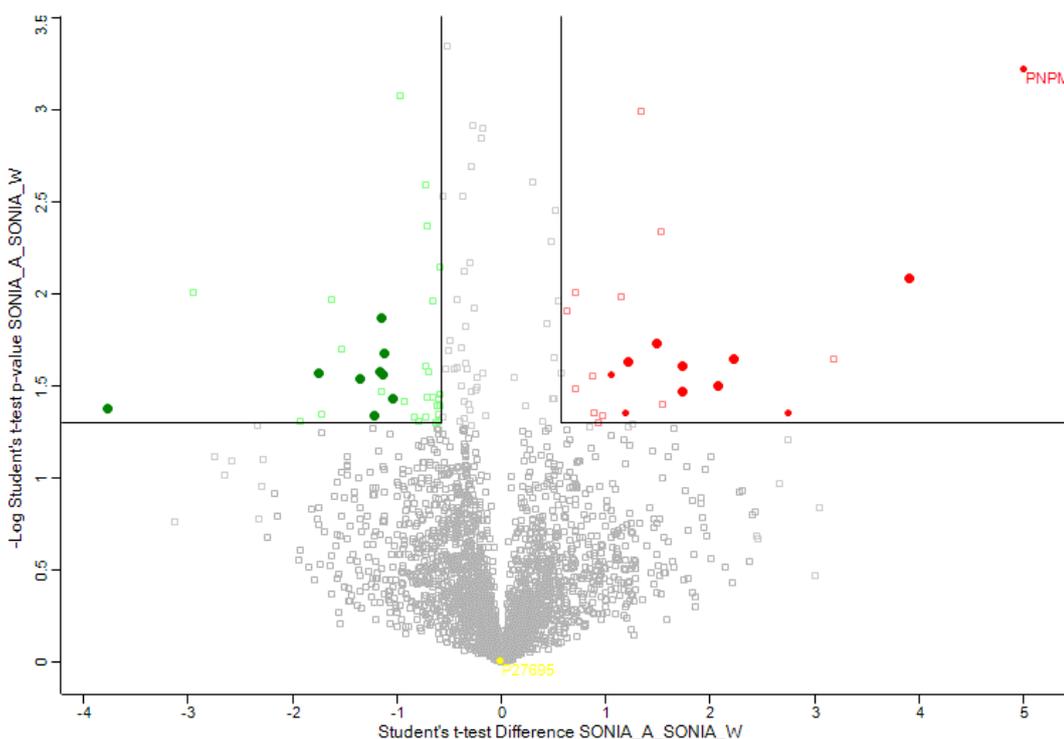


Figura 7. Volcano plot de A vs W. En verde se representan las proteínas significativas en la muestra de proteína salvaje (W) de acuerdo a la combinación de ambos parámetros ($-\log p\text{-value}$ y *t-test difference*) y en rojo se muestran las proteínas más significativas en la muestra de la proteína mutante (A). Las proteínas señaladas en círculos corresponden a aquellas identificadas con más de un péptido (**Tablas 1 y 2**). El punto amarillo corresponde a la proteína APE1. La proteína etiquetada como PNPM es el mutante A.

La comparación de la proteína salvaje frente al mutante (**Figura 7**) nos permite explorar qué implican las mutaciones asociadas a la AML. Por una parte, en cuanto a la pérdida de interacciones por parte de la NPM1, es decir, ligandos “significativamente” enriquecidos en la muestra W y ausentes o muy poco abundantes en A (**Tabla 1**). Entre ellos, destaca la proteína codificada por el gen *CCDC134* (proteína 134 con dominio *coiled-coil*) que tiene un papel en la regulación de la migración e invasión celular en el cáncer gástrico (Zhong *et al*, 2012). Esta proteína podría ser una potencial diana terapéutica del cáncer. También se han encontrado dos polimerasas, POLR2K y POLR1D, aunque identificadas sólo con un péptido (**Tabla AI.1**), relacionadas con reparación, replicación y/o tolerancia al daño en el DNA. Puesto que las polimerasas pueden ayudar a las células cancerosas a tolerar daños del DNA, podrían ser objetivos viables para estrategias terapéuticas (Lange *et al*, 2011). Además, entre las proteínas enriquecidas en la proteína salvaje se encuentran varias proteínas ribosomales como la proteína ribosomal 60S L21 (**Tabla AI.1**), lo cual sugiere que en las células de la AML podría estar desregulada la biosíntesis de ribosomas.

Tabla 1. Proteínas enriquecidas en la proteína salvaje (W) en la comparación con el mutante A. Se representan las proteínas más significativas de acuerdo a los dos parámetros que han sido identificadas por más de un péptido.

Gen	Nombre de la proteína	Péptidos (únicos)	- Log Student's t-test p-value	Student's t-test difference (fold change)
ACADS	Short-chain specific acyl-CoA dehydrogenase, mitochondrial	2 (2)	1,375	-3,770
NDUFA11	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 11	2 (2)	1,569	-1,756
AK6	Adenylate kinase isoenzyme 6	2 (2)	1,541	-1,352
AAR2	Protein AAR2 homolog	2 (2)	1,341	-1,214
CCDC134	Coiled-coil domain-containing protein 134	2 (2)	1,582	-1,167
DCUN1D5	DCN1-like protein 5	8 (8)	1,869	-1,146
GRPEL1	GrpE protein homolog 1, mitochondrial	5 (5)	1,560	-1,139
TXNDC17	Thioredoxin domain-containing protein 17	3 (3)	1,676	-1,125
GPD1L	Glycerol-3-phosphate dehydrogenase 1-like protein	6 (6)	1,435	-1,040

Por otra parte, el mutante A adquiere nuevas interacciones (**Tabla 2 y Tabla AI.2**). Aparece una proteína asociada a quinasa que interviene en la fase S del ciclo celular (SKP1) de la que se han encontrado evidencias que apoyan su papel en la correcta segregación cromosómica (Piva *et al*, 2002).

También se ha encontrado enriquecida la proteína de unión al RNA, FUS, que interviene en el metabolismo del RNA y la regulación de la transcripción. FUS ha sido relacionada con la AML cuando se encuentra asociada a ERG, un regulador transcripcional, y se da una regulación aberrante de la vía de señalización del ácido retinoico, inhibiendo la diferenciación a lo largo del linaje mieloide (Sotoca *et al*, 2016). En el caso de la AML podría estar también alterada dicha vía de señalización.

Además de ellas, también se ha identificado como una proteína sobreexpresada en el mutante, la deacilasa SIRT5. Ésta es una enzima multifuncional que regula varias vías metabólicas vitales en la mitocondria y el citosol. Un estudio llevado a cabo en la universidad de Utah (Deininger), evidenció que la supresión del gen *SIRT5* causaba la inhibición del crecimiento y muerte de células tumorales, mientras que las células sanas no se veían afectadas, lo cual sugiere que esta proteína es clave para la supervivencia de las células cancerosas en la AML y podría ser una diana terapéutica para tratar éste y otros tipos de cáncer mediante el desarrollo de inhibidores de SIRT5.

Tabla 2. Proteínas enriquecidas en el mutante (A) en la comparación con la proteína salvaje. Se representan las proteínas más significativas de acuerdo a los dos parámetros que han sido identificadas por más de un péptido; a excepción de la primera entrada que corresponde al mutante A (NPMmutA) que se diferencia únicamente por un péptido de la proteína salvaje.

Gen	Nombre de la proteína	Péptidos (únicos)	- Log Student's t-test p-value	Student's t-test difference (fold change)
NPM1	NPMmutA	28 (1)	3,219	4,998
APOL2	Apolipoprotein L2	2 (2)	2,085	3,906
SKP1	S-phase kinase-associated prot 1	4 (4)	1,356	2,751
MCU	Calcium uniporter protein, mitoc.	2 (2)	1,643	2,233
TTC27	Tetratricopeptide repeat prot 27	2 (2)	1,498	2,074
PANK4	Pantothenate kinase 4	3 (3)	1,607	1,729
DCTD	Deoxycytidylate deaminase	3 (3)	1,473	1,729
ALKBH5	RNA demethylase ALKBH5	2 (2)	1,729	1,493
SIRT5	NAD-dependent protein deacylase sirtuin-5, mitochondrial	2 (2)	1,631	1,212
FUS	RNA-binding protein FUS	8 (6)	1,355	1,188
EXOSC7	Exosome complex component RRP42	3 (3)	1,562	1,054

Como ya se ha mencionado, la localización aberrante de la NPM1 en la eucemia AML interfiere probablemente con sus funciones, pudiendo estar relacionada con el origen de la enfermedad. Una de las funciones que suscita más interés es su papel en diversos procesos de reparación del DNA. Dicha reparación puede verse afectada en condiciones patológicas. Por tanto, un punto de partida interesante para intentar comprender su implicación en la reparación del DNA y las consecuencias de su deslocalización, es llevar a cabo un análisis proteómico comparativo, como el expuesto en este trabajo, para identificar *hits* implicados en rutas de reparación del DNA.

Además de las anteriormente presentadas, en un estudio similar realizado previamente por el grupo (datos no publicados), se hallaron varias proteínas cuya función se relacionaba con la respuesta y reparación del DNA ante daños. Una de esas proteínas es la APE1, una de las endonucleasas principales en las vías de reparación por escisión de bases (BER), actuando sobre lesiones oxidativas del DNA. La interacción entre ambas proteínas ya ha sido reseñada en trabajos anteriores (Vascotto *et al*, 2009).

La localización subcelular y niveles de expresión aberrantes de APE1 en tumores se han relacionado con una mayor agresividad de los tumores, así como con la aparición de quimio y radioresistencia en diversos tipos de cáncer. La interacción con NPM1 modula varias funciones de APE1: promueve la acumulación de ésta en el nucléolo y estimula su actividad endonucleasa (Poletto *et al*, 2015). Por tanto, la interacción entre APE1 y NPMc+ que parece darse en el citoplasma de las células tumorales (Vascotto *et al*, 2014), podría alterar las vías de reparación BER; lo que afectaría tanto a la proliferación celular como a la capacidad de respuesta a genotoxinas.

Mencionar también que en el análisis se ha obtenido un ligando común a W y A, comparando las interacciones encontradas en ambas muestras respecto al control YFP. Se trata de la proteína *COMM domain containing protein 9*, un posible regulador del complejo cullin-RING E3 ligasa y de la activación de NFκβ.

4.2. EXPERIMENTO F2H

Para profundizar en la interacción NPM1/APE1, quisimos comprobar su interacción *in vivo*, cuando ambas proteínas se coexpresan en células, utilizando una novedosa metodología, F2H, desarrollada recientemente (ChromoTek).

A la vista de los resultados (**Figura 8**), no parece haberse dado interacción entre NPM1 y APE1 ya que no se advierte un punto nítido de fluorescencia roja en el nucléolo co-localizado con el punto verde, que sí aparece en la muestra control. En su lugar se observa una fluorescencia roja dispersa. Esto indica que, en nuestras condiciones, no se detecta interacción. Sin embargo, la técnica es muy reciente y se ha probado en muy pocos sistemas hasta el momento, por lo que no sabemos qué limitaciones puede tener.

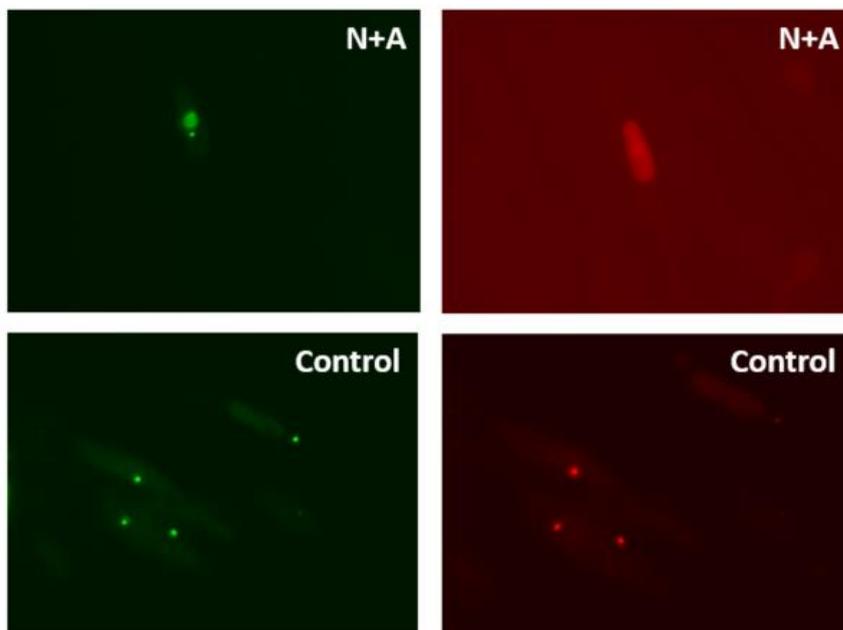


Figura 8. Resultado del experimento F2H. Células F2H co-transfectadas con YFP-NPM1 y con APE1-Cherry, en las que se observa la formación del “punto” verde, pero que, a diferencia del control positivo, no coincide con la tinción roja, indicando que no se detecta interacción entre las dos proteínas.

5. CONCLUSIONES

El estudio proteómico comparativo de la NPM1 presentado en este trabajo es un punto de partida para conocer mejor los mecanismos desencadenantes de la AML. Permite realizar un análisis masivo para la identificación de los ligandos de la NPM1 salvaje y la variante mutada, mutante A, que es la proteína más frecuentemente mutada.

Gracias al análisis se han podido identificar varios ligandos enriquecidos en el mutante respecto a la NPM1 salvaje y viceversa. Esto sugiere que la proteína alterada NPM1 que queda aberrantemente deslocalizada en el citoplasma, pierde la capacidad de unión a ciertos ligandos y/o puede unir nuevos ligandos respecto a la variante salvaje

afectando directamente a su correcto funcionamiento e intervención en la regulación de procesos de reparación de daños del DNA

Concretamente, el mutante A presenta ligandos significativamente enriquecidos respecto a la proteína salvaje implicados en la proliferación celular y relacionadas con leucemia y otros tipos de cáncer como FUS, SIRT5 y SKP1. Por su parte, también se han identificado ligandos propios de la NPM1 que parecen perderse o unirse menos, al mutante. Destacan las polimerasas POLR1D y POLR2K que cumplen un papel de respuesta y reparación de los daños del DNA.

Por último, la interacción de la proteína mutante con ligandos clave de la NPM1 y su reclutamiento en el citoplasma puede desencadenar la transformación de las células, al verse inhibidas sus funciones de reparación y/o estimulada la proliferación anómala por carecer de control en el procesamiento del DNA, como sugiere el enriquecimiento de la proteína SKP1 en la NPM1 salvaje frente al mutante A.

Por otro lado, a pesar de que los resultados referentes a la interacción de NPM1 y APE1 no nos llevan a poder afirmar la interacción entre ambas proteínas, es interesante el estudio de esta u otras interacciones con proteínas implicadas en las vías de reparación BER. Los tratamientos actuales para el cáncer se basan en una terapia combinada de inhibidores específicos de la reparación del DNA con fármacos dañinos para el DNA con el fin de conseguir una mayor sensibilidad de las células cancerosas a los agentes quimioterapéuticos. Recordemos que la sobreexpresión de proteínas de las vías BER se han relacionado con quimio y radio-resistencia, en concreto APE1 ha sido relacionada con resistencia a diversos quimioterapéuticos como *cisplatin* o *bleomycin* así como a la radiación ionizante (Robertson *et al*, 2001).

Por tanto, los resultados del análisis proteómico revelan varias proteínas que podrían resultar interesantes como dianas para el tratamiento de la AML.

6. BIBLIOGRAFÍA

Artículos:

Arregi I, Falces J, Olazabal-Herrero A, Alonso-Mariño M, Taneva SG, Rodríguez JA, Urbaneja MA. y Bañuelos S. (2015). Leukemia-Associated Mutations in

Nucleophosmin Alter Recognition by CRM1: Molecular Basis of Aberrant Transport. PLoS One (Public Library of Science) 10(6): e0130610.

Box JK, Paquet N, Adams MN, Boucher D, Bolderson E, O'Byrne KJ. y Richard DJ. (2016). Nucleophosmin. From structure and function to disease development. BMC Molecular Biology 17:19.

Cox J. y Mann M. (2008). MaxQuant enables high peptide identification rates, individualied p.p.b.-range mass accuracies and proteome-wide protein quantification. Natura Biotechnology 26, 1367-1372.

De Kouchkovsky I. y Abdul-Hay M. (2016). Acute myeloid leukemia: a comprehensive review and 2016 update. Blood Cancer Journal 6(7): e441.

Deininger MW. (2016-actualidad). SIRT5 as a Therapeutic Target in Acute Myeloid Lekemia. National Institutes of Health, University of Utah (EEUU).

Falini B, Martelli MP, Bolli N, Sportoletti P, Liso A, Tiacci E. y Heferlach T. (2011). Acute myeloid leukemia with mutated nucleophosmin (*Npm1*): is it a distinct entity? Blood Journal. 117:1109-1120.

Falini B, Nicoletti I, Martelli MF. y Mecucci C. (2007). Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): biologic and clinical features. Blood Journal. 109:874-885.

Federici L, Acrovito A, Scaglione GL, Scalon F, Lo Sterzo C, Di Matteo A, Falini B, Giardina B. y Brunori M. (2010). Nuclephosmin C-Terminal Leukemia-Associated Domain Interacts With G-Rich Quadruplex Forming DNA. The Journal of Biological Chemistry 285(48): 37138-37149.

Federici L. y Falini B. (2013). Nucleophosmin mutations in acute myeloid leukemia: A tale of proteín unfolding and mislocalization. Protein Science 22(5): 545-556.

Grummit CG, Townsley FM, Johnson CM, Warren AJ. y Bycroft M. (2008). Structural Consequences of Nucleophosmin Mutations in Acute Myeloid Leukemia. The Journal of Biological Chemistry 283(34): 23326-23332.

Koike A, Nishikawa H, Wu W, Okada Y, Venkitaraman AR. y Ohta T. (2010). Recruitment of Phosphorylated NPM1 to Sites of DNA Damage through RNF8-Dependent Conjugates. *Cancer Research Journal* 70 (17), 6746-6756.

Konoplev S, Huang X, Drabkin HA, Koeppen H, Jones D, Kantarijan HM, García-Manero G, Chen W, Weina LJ. y Bueso-Ramos CE. (2009). Cytoplasmic Localization of Nucleophosmin in Bone Marrow Blasts of Acute Myeloid Leukemia Patients is not Completely Concordant with NPM1 Mutation and is not Predictive of Prognosis. *Cancer Journal* 115 (20): 4737-4744.

Lagunas-Rangel FA. (2016). Leucemia mieloide aguda: una perspectiva de los mecanismos moleculares del cáncer. *Gaceta Mexicana de Oncología* 15 (3): 150-157.

Lange SS, Takata K. y Wood RD. (2011). DNA polymerases and cancer. *Nature Reviews Cancer* 11(2): 96-110.

Lee HH, Kim HS, Kang JY, Lee BI, Ha JY, Yoon HJ, Lim SO, Jung G, Suh SW. (2007). Crystal structure of human nucleophosmin-core reveals plasticity of the pentamer-pentamer interface. *Proteins* 69: 672-678.

Löwenber B, Downing JR. y Burnett A. (1999). Acute Myeloid Leukemia. *The New England Journal of Medicine* 341: 1051-1062.

Piva R, Liu J, Chiarle R, Podda A, Pagano M. y Inghirami G. (2002). In vivo interference with SKP1 function leads to genetic instability and neoplastic transformation. *Molecular and Cellular Biology*. Vol 22, No 23 p. 8375-8387.

Poletto M, Mafatti MC, Dorjsuren D, Scognamiglio PL, Marasco D, Vascotto C, Jadhav A, Maloney DJ, Wilson III DM, Simeonov A. y Tell G. (2015). Inhibitors of the apurinic/aprimidinic endonuclease 1 (APE1)/nucleophosmin (NPM1) interaction that display anti-tumor properties. *Molecular Carcinogenesis Journal*. 55:688-704.

Robertson KA, Bullock HA, Xu Yi, Tritt Renee, Zommerman TM, Foster RS, Einhorn LH y Kelley MR. (2001). Altered Expression of APE1/ref-1 in Germ Cell Tumors and Overexpression in NT2 Cells Confers Resistance to Bleomycin and Radiation. *Tumor Biology* 61 (5): 2220-2225.

Scott DD. y Oeffinger M. (2016). Nucleolin and Nucleophosmin: nucleolar proteins with multiple functions in DNA repair. *Biochemistry and Cell Biology* 94(5):419-432.

Sotoca AM, Prange KHM, Reijnders B, Mandoli A, Nguyen LN, Stunnenberg HG. Y Martens JHA. (2016). The oncofusion protein FUS-ERG targets key hematopoietic regulators and modulates the all-trans retinoic acid signaling pathway in t (16;21) acute myeloid leukemia. *Oncogene* 35, 1965-1976.

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M y Cox J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* 13, 731-740.

Rothbauer U, Leonhardt H., Zolghadr K. y Mortusewicz O. (2010). Fluorescent two-hybrid (F2H) assay for direct visualization of protein interactions in living cells. *US* 12/811,712. 11/11/2010. Universidad Ludwig-Maximilians, Munich, Alemania.

Vascotto C, Fantini D, Romanello M, Cesaratto L, Deganuto M, Leonardi A, Radicella JP, Kelley MR, D'Ambrosio C, Scaloni A, Quadrifoglio F. y Tell G. (2009). APE1/Ref-1 interacts with NPM1 within nucleoli and plays a role in the rRNA quality control process. *Molecular Cell Biology* 29(7): 1834-54.

Vascotto C, Lirussi L, Poletto M, Tiribelli M, Damiani D, Fabbro D, Damante G, Demple B Colombo E y Tell G. (2014). Functional regulation of the apurinic/aprimidinic endonuclease 1 by nucleophosmin: impact on tumor biology. *Oncogen* 29;33(22):2876-87.

Zhong J, Zhao M, Luo Q, Ma Y, Liu J, Wang J, Yang M, Yuan X, Sang J. y Huang C. (2012). CCDC134 is down-regulated in gastric cancer and its silencing promotes cell migration and invasion of GES-1 and AGS cells via MAPK pathway. *Molecular Cell Biochemistry* 372: 1-8.

Páginas web:

American Cancer Society. 17 de mayo de 2017. Recuperado de: <https://www.cancer.org>.

ANEXO

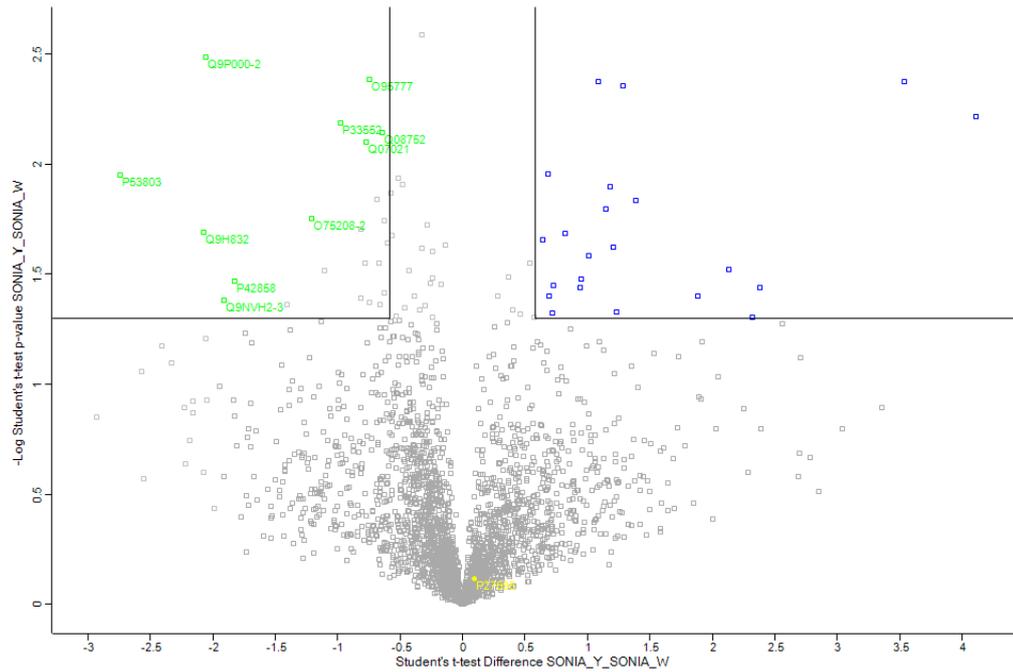


Figura 1. Volcano plot de Y vs W. En verde las proteínas significativas en la muestra de proteína salvaje (W) de acuerdo a la combinación de ambos parámetros ($-\log p\text{-value}$ y $t\text{-test difference}$). En azul las proteínas más significativas en la muestra de la etiqueta (Y). El punto amarillo corresponde a la proteína APE1.

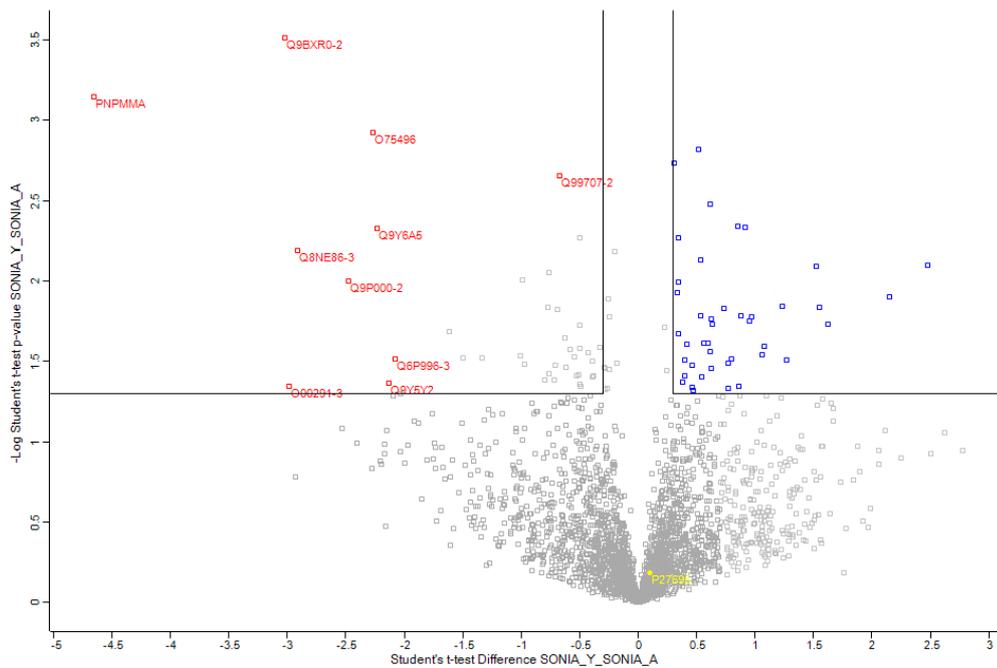


Figura 2. Volcano plot de Y vs A. En rojo las proteínas significativas en la muestra de proteína mutante (A) de acuerdo a la combinación de ambos parámetros ($-\log p\text{-value}$ y $t\text{-test difference}$). En azul las proteínas más significativas en la muestra de la etiqueta (Y). El punto amarillo corresponde a la proteína APE1.

Tabla 1. Proteínas enriquecidas en la proteína salvaje (W) en la comparación con el mutante A. Se representan las proteínas más significativas de acuerdo a los dos parámetros. En gris se muestran las proteínas identificadas por un solo péptido.

Gen	Nombre de la proteína	Péptidos (únicos)	- Log Student's t-test p-value	Student's t-test difference (fold change)
ACADS	Short-chain specific acyl-CoA dehydrogenase, mitochondrial	2 (2)	1,375	-3,770
POLR2K	DNA-directed RNA polymerases I, II, and III subunit RPABC4	1 (1)	2,012	-2,946
NDUFA11	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 11	2 (2)	1,569	-1,756
POLR1D	DNA-directed RNA polymerases I and III subunit RPAC2	1 (1)	1,347	-1,725
SYPL1	Synaptophysin-like protein 1	1 (1)	1,971	-1,630
MRPL53	39S ribosomal protein L53, mitochondrial	1 (1)	1,699	-1,535
AK6	Adenylate kinase isoenzyme 6	2 (2)	1,541	-1,352
AAR2	Protein AAR2 homolog	2 (2)	1,341	-1,214
CCDC134	Coiled-coil domain-containing protein 134	2 (2)	1,582	-1,167
DCUN1D5	DCN1-like protein 5	8 (8)	1,869	-1,146
GPN1	GPN-loop GTPase 1	1 (1)	1,469	-1,145
GRPEL1	GrpE protein homolog 1, mitochondrial	5 (5)	1,560	-1,139
TXNDC17	Thioredoxin domain-containing protein 17	3 (3)	1,676	-1,125
GPD1L	Glycerol-3-phosphate dehydrogenase 1-like protein	6 (6)	1,435	-1,040
PPP6R1	Serine/threonine-protein phosphatase 6 regulatory subunit 1	4 (4)	3,078	-0,973
XPO7	Exportin-7	9 (9)	1,417	-0,933
MRRF	Ribosome-recycling factor, mitochondrial	2 (2)	1,335	-0,837
C7ORF50	Uncharacterized protein C7orf50	4 (4)	1,429	-0,824
GLOD4	Glyoxalase domain-containing protein 4	10 (10)	1,330	-0,728
BOLA2	BolA-like protein 2	5 (5)	2,592	-0,727

Continuación tabla 1. Proteínas enriquecidas en la proteína salvaje (W) en la comparación con el mutante A. Se representan las proteínas más significativas de acuerdo a los dos parámetros. En gris se muestran las proteínas identificadas por un solo péptido.

Gen	Nombre de la proteína	Péptidos (únicos)	- Log student's t-test p-value	Student's t-test difference (fold change)
AUP1	Ancient ubiquitous protein 1	3 (3)	1,608	-0,725
SSR1	Translocon-associated protein subunit alpha	3 (3)	1,441	-0,708
COQ9	Ubiquinone biosynthesis protein COQ9, mitochondrial	2 (2)	2,367	-0,705
ZC3H4	Zinc finger CCCH domain-containing protein 4	4 (4)	1,577	-0,703
LTV1	Protein LTV1 homolog	4 (4)	1,442	-0,663
RPL21	60S ribosomal protein L21	9 (9)	1,964	-0,654
AIP	AH receptor-interacting protein	7 (7)	1,396	-0,609
NPM3	Nucleoplasmin-3	3 (3)	1,433	-0,606
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	21 (21)	1,350	-0,604
ACADM	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	9 (9)	1,394	-0,587
KDELC2	KDEL motif-containing protein 2	2 (2)	2,145	-0,582
WDR1	WD repeat-containing protein 1	6 (6)	1,456	-0,581

Tabla 2. Proteínas enriquecidas en el mutante A (A) en la comparación con la proteína salvaje. En gris se marcan aquellas proteínas identificadas por un único péptido, y que por tanto no se consideran significativas de la NPM1 salvaje. La primera entrada corresponde al mutante A (NPMmutA) que se diferencia únicamente por un péptido de la proteína salvaje.

Gen	Nombre de la proteína	Péptidos (únicos)	- Log Student's t-test p-value	Student's t-test difference (fold change)
NPM1	NPMmutA	28 (1)	3,219	4,998
APOL2	Apolipoprotein L2	2 (2)	2,085	3,906
PMF1	Polyamine-modulated factor 1	1 (1)	1,646	3,185
SKP1	S-phase kinase-associated protein 1	4 (3)	1,356	2,751
MCU	Calcium uniporter protein, mitochondrial	2 (2)	1,643	2,233

Continuación Tabla 2. Proteínas enriquecidas en el mutante A (A) en la comparación con la proteína salvaje. En gris se marcan aquellas proteínas identificadas por un único péptido, y que por tanto no se consideran significativas de la NPM1 salvaje.

Gen	Nombre de la proteína	Péptidos (únicos)	- Log student's t-test p-value	Student's t-test difference (fold change)
TTC27	Tetratricopeptide repeat protein 27	2 (2)	1,498	2,074
PANK4	Pantothenate kinase 4	3 (3)	1,607	1,729
DCTD	Deoxycytidylate deaminase	3 (3)	1,473	1,729
SDF2L1	Stromal cell-derived factor 2-like protein 1	1 (1)	1,398	1,542
CPT1A	Carnitine O-palmitoyltransferase 1, liver isoform	1 (1)	2,335	1,527
ALKBH5	RNA demethylase ALKBH5	2 (2)	1,729	1,493
EIF2D	Eukaryotic translation initiation factor 2D	1 (1)	2,990	1,333
SIRT5	NAD-dependent protein deacylase sirtuin-5, mitochondrial	2 (2)	1,631	1,212
FUS	RNA-binding protein FUS	8 (6)	1,355	1,188
HSD17B8	Estradiol 17-beta-dehydrogenase 8	1 (1)	1,987	1,140
EXOSC7	Exosome complex component RRP42	3 (3)	1,562	1,054
SMYD5	SET and MYND domain-containing protein 5	1 (1)	1,336	0,963
PUM1	Pumilio homolog 1	12 (1)	1,358	0,891
HNRNPLL	Heterogeneous nuclear ribonucleoprotein L-like	4 (2)	1,555	0,879
RPS17	40S ribosomal protein S17	9 (9)	2,005	0,713
RALY	RNA-binding protein Raly	7 (7)	1,488	0,712
PPP1R7	Protein phosphatase 1 regulatory subunit 7	3 (3)	1,910	0,626