



Trabajo Fin de Grado/ Gradu Amaierako Lana

Grado en Bioquímica y Biología Molecular/ Biokimika eta Biologia Molekularra Gradua

Búsqueda de señales de transporte nucleocitoplásmico en la proteína KMT2D

Autor/a/ Egilea:

Andrea Jáñez Pedrayes

Director/a/ Zuzendaria:

Dr. Jose Antonio Rodríguez Pérez

ÍNDICE

1. INTRODUCCIÓN Y OBJETIVOS	1
2. MATERIALES Y MÉTODOS	3
2.1. Análisis bioinformático	3
2.1.1. Predicción de señales de localización nuclear	3
2.1.2. Predicción de señales de exportación nuclear.....	3
2.1.3. Búsqueda de mutaciones clínicamente relevantes en las posibles NLSs/NESs.....	3
2.2. Análisis experimental de las señales de transporte nucleocitoplásmico.....	4
2.2.1. Análisis experimental de potenciales NLSs.....	4
2.2.2. Análisis experimental de potenciales NESs: Ensayo de exportación nuclear	4
2.2.3. Clonación (generación de plásmidos)	5
2.2.4. Cultivo, transfección, tratamiento y fijación de las células	6
2.2.5. Microscopía de fluorescencia	7
3. RESULTADOS	7
3.1. Predicción y selección de potenciales señales de transporte nucleocitoplásmico en la proteína KMT2D mediante análisis bioinformático	7
3.1.1. Selección de una potencial NLS en KMT2D.....	7
3.1.2. Selección de dos potenciales NESs en KMT2D.....	8
3.1.3. Una mutación clínicamente relevante en KMT2D afecta a una potencial NES	8
3.2. Resultados de la clonación	9
3.2.1. Transformación de las bacterias	9
3.2.2. Análisis mediante PCR y electroforesis de los insertos presentes en los plásmidos	9
3.2.3. Secuenciación de los insertos	10
3.3. La región 1680- 1708 de la proteína KMT2D es una señal de localización nuclear funcional ..	10
3.4. La región 4549- 4567 de la proteína KMT2D funciona como una señal de exportación nuclear débil.....	11
4. DISCUSIÓN	13
5. CONCLUSIONES Y FUTUROS ESTUDIOS	14
6. BIBLIOGRAFÍA	14

1. INTRODUCCIÓN Y OBJETIVOS

El síndrome de Kabuki es una enfermedad congénita rara que afecta aproximadamente a 1 de cada 32.000 nacimientos. Se caracteriza principalmente por la presencia de rasgos faciales dismórficos distintivos (que recuerdan al maquillaje de los actores del teatro japonés Kabuki), retraso en el desarrollo y crecimiento postnatal, discapacidades intelectuales y anomalías esqueléticas (Cheon y Ko, 2015). Muchos pacientes con este síndrome sufren defectos cardíacos congénitos, como dilatación de la aorta y defectos septales ventriculares (Digilio *et al.*, 2017). Además, recientemente se ha sugerido que estos pacientes también tienen una predisposición a padecer cáncer (Teranishi *et al.*, 2018).

Este espectro de manifestaciones clínicas tan amplio se debe a la pérdida de la función del gen KMT2D (anteriormente denominado MLL2), que se ha encontrado mutado en hasta un 80% de los casos de Kabuki (Cheon y Ko, 2015). La mayoría de mutaciones en este gen se dan de manera espontánea, aunque también se han descrito mutaciones heredadas con un patrón autosómico dominante. El gen KMT2D se localiza en la región cromosómica 12q13.12 y codifica para la síntesis de la histona-lisina N- metiltransferasa 2D o KMT2D, una proteína de 5537 aminoácidos, que es esencial en el desarrollo embrionario temprano y que se expresa también en numerosos tejidos adultos (Froimchuk *et al.*, 2017).

KMT2D forma un complejo proteico mediante su asociación con WRAD (WDR5, RbBP5, ASH2L, DPY30), UTX, NCOA6, PA1 y PTIP (Figura 1). KMT2D actúa como *scaffold* para el resto de proteínas,

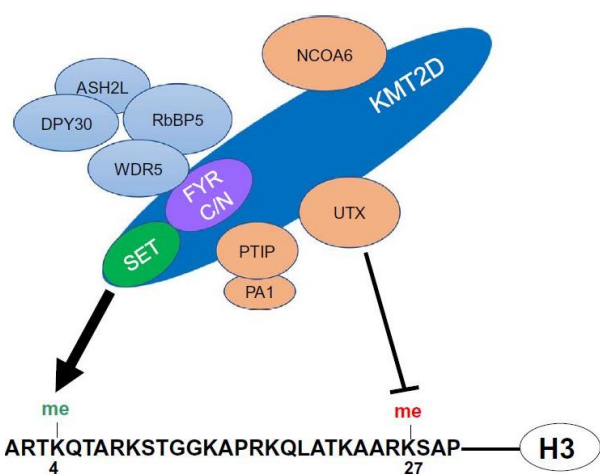


Figura 1. Complejo proteico de KMT2D (Froimchuk *et al.*, 2017).

KMT2D actúa en la regulación epigenética de la transcripción de una gran variedad de genes implicados en la embriogénesis y el desarrollo. Concretamente, actúa selectivamente en regiones potenciadoras (*enhancer regions*) específicas, dependiendo del tipo celular y estado de diferenciación. Modula la accesibilidad de la cromatina, permitiendo el reclutamiento de distintos factores de transcripción y la consiguiente expresión de genes específicos, entre los que se incluyen genes Hox (Guo *et al.*, 2013). Así, KMT2D juega un papel crítico en la regulación del desarrollo, diferenciación y metabolismo, lo que explica que las mutaciones en este gen en la línea germinal sean tan perjudiciales.

Además de las mutaciones en la línea germinal, asociadas al síndrome de Kabuki, se ha visto que las mutaciones somáticas en KMT2D promueven la formación de tumores. Hasta la fecha se han identificado cientos de mutaciones de KMT2D en tumores humanos, siendo esta proteína una de las más frecuentemente mutadas en el cáncer (Rao y Dou, 2015). Se han descrito alteraciones de KMT2D en meduloblastomas, leucemia mieloide aguda, cáncer de pulmón, de intestino, de endometrio, de próstata, hepático y renal, entre otros (Guo *et al.*, 2013; Rao y Dou, 2015). Además, KMT2D está implicada en mecanismos de resistencia terapéutica inducidos por el empleo de inhibidores de la PI3K α en casos de cáncer de mama ER (receptor de estrógenos) -positivo, ya que permite el reclutamiento de ER a lugares específicos de la cromatina para activar la transcripción de genes que favorecen el crecimiento tumoral (Toska *et al.*, 2017). En conjunto, estos estudios demuestran que la proteína KMT2D está íntimamente relacionada con muchas patologías humanas.

KMT2D es una metiltransferasa de histonas que desempeña su función en el núcleo celular. Por ello, tiene que ser transportada desde el citoplasma, donde es sintetizada, al núcleo. Las proteínas con una masa atómica superior a los 40 kDa, como es el caso de KMT2D (600 kDa), deben ser transportadas activamente entre el núcleo y el citoplasma a través de los poros nucleares. Para ello, requieren la acción de receptores de transporte solubles llamados importinas (Imp) y exportinas, que reconocen señales de localización nuclear (NLS) y señales de exportación nuclear (NES) respectivamente, presentes en la secuencia de las proteínas que han de transportar (Cautain *et al.*, 2015).

Las NLSs “clásicas” son secuencias cortas que contienen *clusters* de aminoácidos cargados positivamente, típicamente lisina o arginina. Se subdividen en monopartitas y bipartitas, en función de si están formadas por un único *cluster* o por dos *clusters* separados por una región *linker* de unos 10 aminoácidos. Las NLSs clásicas se unen a la proteína adaptadora Imp- α , que a su vez se une al receptor de transporte Imp- β 1, formando un complejo ternario que es transportado al núcleo. Hay otros tipos de señales NLS “no clásicas”, como la denominada PY-NLS, que consiste en un motivo N-terminal hidrofóbico, un residuo de arginina central y una secuencia PY C-terminal. Las proteínas con PY-NLS son importadas al núcleo por el receptor Imp- β 2 (Cautain *et al.*, 2015).

En cuanto a la exportación nuclear, la NES “rica en leucina”, que es reconocida por el receptor de exportación nuclear CRM1, es la señal que ha sido más caracterizada. Se trata de una secuencia aminoacídica corta con varios residuos hidrofóbicos. Tras los primeros estudios que se realizaron, se propuso una secuencia NES consenso [ϕ^1 - (X)₂₋₃ - ϕ^2 - (X)₂₋₃ - ϕ^3 - (X) - ϕ^4] con cuatro residuos hidrofóbicos conservados (representados por ϕ ; pueden ser L, V, I, F o M) separados por un número variable de aminoácidos (representados por X), preferentemente cargados, polares o de pequeño tamaño. Sin embargo, este patrón consenso se ha ido refinando a medida que se han identificado y confirmado experimentalmente más NESs (Cautain *et al.*, 2015; Prieto *et al.*, 2014). Además, estudios estructurales han revelado que CRM1 contiene 5 bolsillos de unión para los residuos hidrofóbicos conservados de las

NESs, por lo que en base a esto se ha propuesto una NES consenso con una posición hidrofóbica adicional (Güttler *et al.* 2010).

Cabe señalar, por otra parte, que algunas proteínas pueden ser transportadas dentro y fuera del núcleo a través de la interacción con otras proteínas que contengan señales de transporte funcionales, a pesar de que ellas no tengan ni NLS ni NES (Cautain *et al.*, 2015).

En vista de la gran relevancia tanto fisiológica como patológica de KMT2D, con este proyecto se va a tratar de ahondar en la caracterización de esta proteína. Dada su función en el núcleo, partimos de la hipótesis de que KMT2D puede tener señales de transporte nucleocitoplásmico funcionales. Por tanto, el objetivo principal del trabajo va a ser la identificación de señales NLS y NES en la secuencia de KMT2D para conocer los mecanismos de transporte empleados por esta proteína. Además, otro de los objetivos va a ser el estudio de mutaciones patológicas en las potenciales señales de transporte, con el fin de analizar si la patología puede deberse al transporte alterado de KMT2D.

2. MATERIALES Y MÉTODOS

2.1. Análisis bioinformático

2.1.1. Predicción de señales de localización nuclear

Para la búsqueda de NLSs en la secuencia de la proteína KMT2D se emplearon tres programas web: Wregex (Prieto *et al.*, 2014), PSORT II (una versión actualizada de PSORT (Nakai y Horton, 1999)), y ELM (Dinkel *et al.*, 2016). Estos programas determinan la secuencia y localización de las potenciales NLSs aplicando diferentes algoritmos. La secuencia aminoacídica de KMT2D en formato FASTA que se introdujo en los programas se obtuvo de la base de datos Uniprot.

2.1.2. Predicción de señales de exportación nuclear

Para la predicción de las NESs también se utilizaron los programas Wregex (Prieto *et al.*, 2014) y ELM (Dinkel *et al.*, 2016), y además otros dos programas especializados en la predicción de NESs: LocNES (Xu *et al.*, 2015) y NetNES (la Cour *et al.*, 2004). Ambos funcionan de manera similar a los previamente descritos, aunque aplican algoritmos diferentes.

2.1.3. Búsqueda de mutaciones clínicamente relevantes en las posibles NLSs/NESs

Se realizó una búsqueda en la base de datos ClinVar (NCBI) y en el buscador genómico Ensembl de las mutaciones registradas en la secuencia de la proteína KMT2D, la mayoría de las cuales están relacionadas con el Síndrome de Kabuki. Concretamente, se buscaron aquellas mutaciones en las secuencias de las potenciales señales de importación y exportación nuclear seleccionadas a partir de los programas de predicción descritos más arriba.

2.2. Análisis experimental de las señales de transporte nucleocitoplásmico

2.2.1. Análisis experimental de potenciales NLSs

Para determinar la funcionalidad de la NLS candidata (cNLS) la señal se insertó en el plásmido pEYFP-C1, detrás de la secuencia que codifica para la síntesis de la proteína fluorescente YFP. De este modo, la expresión del plásmido da lugar a YFP fusionada a la cNLS (YFP-cNLS). Los plásmidos fueron transfectados en células HeLa y células embrionarias de riñón humano 293 (293T) y se estudió la localización subcelular de la proteína YFP-cNLS por microscopía de fluorescencia. Si la cNLS es funcional, la proteína YFP-cNLS se localizará mayoritariamente en el núcleo celular. En cambio, si no es una verdadera NLS se observará fluorescencia en el citoplasma y en el núcleo, la localización normal de la proteína YFP (Figura 2).

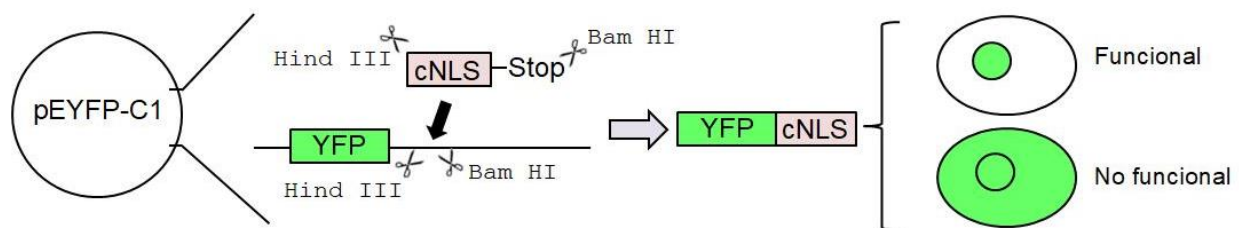


Figura 2. Análisis experimental de potenciales NLSs. La NLS en estudio (cNLS) se clona en el plásmido pEYFP-C1 detrás de la secuencia de la proteína YFP, mediante las endonucleasas Hind III y Bam HI. Tras la transfección de los plásmidos en las células, se estudia la actividad de la cNLS mediante el análisis de la distribución subcelular de la proteína YFP-cNLS. Si la proteína fluorescente se localiza mayoritariamente en el núcleo la cNLS se considera funcional, mientras que si se localiza tanto en el núcleo como en el citoplasma significa que no se trata de una verdadera NLS.

2.2.2. Análisis experimental de potenciales NESs: Ensayo de exportación nuclear

Para el estudio de las NESs candidatas (cNES) se empleó un ensayo de exportación nuclear desarrollado por Henderson y Eleftheriou (Henderson y Eleftheriou, 2000). En este ensayo se emplea el plásmido Rev 1.4-GFP, que codifica una proteína de fusión formada por la proteína Rev 1.4 del virus de la inmunodeficiencia humana (que contiene una NLS fuerte y una NES mutada no funcional) y la proteína fluorescente verde GFP. La señal en estudio se inserta entre las secuencias de Rev 1.4 y GFP, por lo que la expresión del plásmido da lugar a la proteína quimérica fluorescente Rev 1.4-cNES-GFP.

Los plásmidos fueron transfectados en células HeLa y posteriormente las células fueron tratadas solo con cicloheximida (CHX; Sigma) o con CHX y actinomicina D (Act. D; Sigma). La CHX es un inhibidor de la síntesis proteica, de modo que incubar las células con este compuesto asegura que la fluorescencia que se detecte en el citoplasma corresponde a proteínas que han sido exportadas del núcleo gracias a la cNES y no a proteínas que se acaben de sintetizar. La Act. D es un inhibidor de la importación nuclear —mediada por la NLS de Rev—, por lo que tras su adición todas las proteínas que hayan sido exportadas por la cNES se quedan en el citoplasma, lo que facilita la identificación de las NESs débiles. Tras el tratamiento de las

células se procede al estudio de la distribución subcelular de las proteínas por microscopía de fluorescencia. Si la cNES funciona, las proteínas se localizarán en el citoplasma debido a la actividad exportadora de la señal. En cambio, si la cNES no es funcional, la fluorescencia se observará en el núcleo celular (Figura 3).

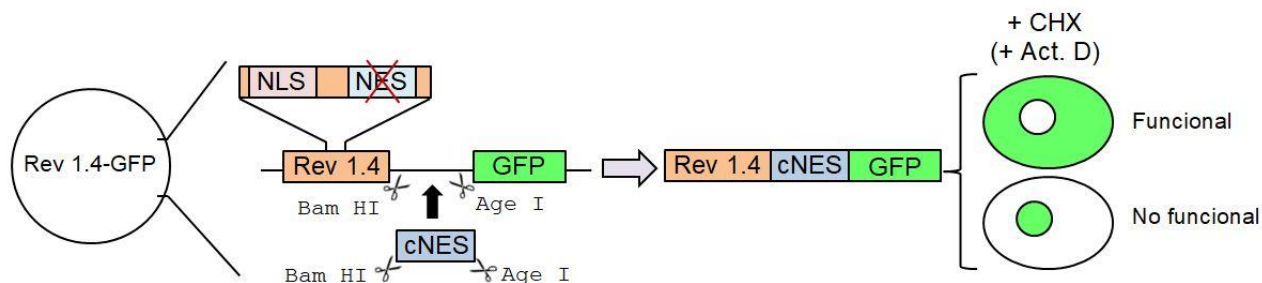


Figura 3. Análisis experimental de potenciales NESs: Ensayo de exportación nuclear. La NES en estudio (cNES) se clona mediante las endonucleasas Bam HI y Age I en el plásmido Rev 1.4-GFP, entre las secuencias de las proteínas Rev 1.4 y GFP. Tras la transfección de los plásmidos y el correspondiente tratamiento de las células solo con CHX o con CHX y Act D, se estudia la actividad de la cNES mediante el análisis de la distribución subcelular de la proteína quimérica Rev 1.4-cNES-GFP. Si la proteína fluorescente se localiza en el citoplasma la cNES se considera funcional, mientras que si se localiza en el núcleo significa que no se trata de una verdadera NES.

2.2.3. Clonación (generación de plásmidos)

Para clonar la cNLS y las cNESs se usaron gBlocks. Los gBlocks son fragmentos de DNA de doble hebra comerciales, sintetizados por la empresa IDT (Integrated DNA Technologies) según la secuencia enviada por el investigador. Se diseñó una secuencia de DNA de 340 pb que contenía las secuencias nucleotídicas que codifican las cuatro señales en estudio— cNLS, cNES 1, cNES 2 y cNES 2 mutada—flanqueadas por las secuencias de corte reconocidas por las correspondientes endonucleasas de restricción (Figura 4).



Figura 4. Diseño de la secuencia del gBlock. Se incluyeron las secuencias de las cuatro señales en estudio (cNLS, cNES 1, cNES 2 y cNES 2 mutada) flanqueadas por sus respectivas secuencias de corte para las endonucleasas y nucleótidos aleatorios a modo de separación para facilitar la digestión. En la parte inferior de la imagen se indican las endonucleasas específicas empleadas para cada tipo de señal junto con sus secuencias de corte.

El DNA del gBlock se digirió con las endonucleasas de restricción específicas Bam HI, Age I y Hind III, obteniendo así los distintos insertos con extremos cohesivos, que posteriormente se purificaron empleando el QIAquick PCR Purification Kit de Qiagen y se eluyeron en agua. Los insertos se ligaron mediante la acción de la DNA ligasa con los vectores (Rev 1.4-GFP o pEYFP-C1), que se encontraban previamente cortados con las enzimas correspondientes y desfosforilados. Se prepararon 2 mezclas de ligación para cada tipo de vector, una de ellas con inserto y la otra sin inserto, para utilizarla como control negativo.

Se transformaron bacterias de la cepa *E.coli* DH5 α con los plásmidos con y sin inserto mediante el método del choque térmico. Las bacterias se sembraron en placas de Petri con kanamicina, de modo que solo pudiesen crecer aquellas bacterias que hubiesen incorporado el plásmido, que confiere resistencia a dicho antibiótico. Las placas se incubaron *overnight* a 37°C para facilitar el crecimiento de las colonias.

Para comprobar si los plásmidos habían incorporado el inserto correctamente se realizó una PCR y posterior electroforesis de cinco muestras de colonias para cada tipo de plásmido. Para la PCR se emplearon un par de cebadores específicos para cada tipo de vector, que amplifican un fragmento del plásmido en el que está incluida la región donde teóricamente se ha insertado la cNLS/cNES. Una vez amplificados los plásmidos, se realizó una electroforesis en un gel de agarosa con el objetivo de determinar si los plásmidos contenían un solo inserto, más de uno o ninguno, en función del tamaño del producto amplificado (Figura 5). En base a los resultados de la electroforesis se eligieron varias colonias de cada tipo y se procedió a la purificación de sus plásmidos con el kit Miniprep de Qiagen. Por último, las muestras se enviaron a secuenciar a la empresa Stab Vida para identificar las señales insertadas en cada una de ellas y comprobar que la inserción se había producido correctamente.

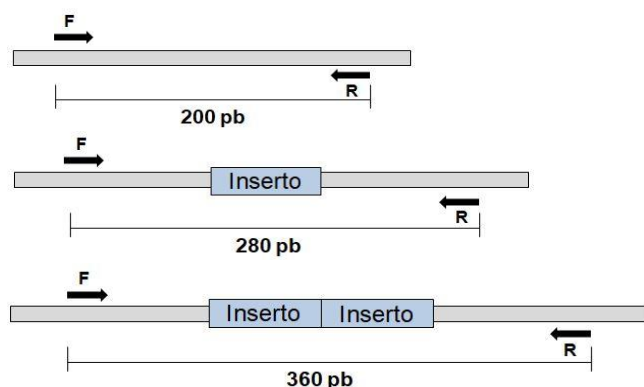


Figura 5. Representación esquemática de la región de los plásmidos que se amplificó por PCR. Cuando no hay inserto, la zona amplificada tiene un tamaño de 200 pb. Al incorporarse un inserto, el tamaño aumenta aproximadamente 80 pb, por lo que el fragmento amplificado es de 280 pb. En caso de incorporarse dos insertos, el tamaño es en torno a las 360 pb.

2.2.4. Cultivo, transfección, tratamiento y fijación de las células

Las células 293T y las células HeLa fueron cultivadas en medio de cultivo Eagle modificado de Dulbecco (DMEM), suplementado con 10% de suero fetal bovino (FBS), 100 U/ml penicilina y 100 μ g/ml estreptomycin (P/S) (todo ello de Invitrogen). Veinticuatro horas antes de la transfección las células se sembraron en una placa de 12 pocillos con un cubreobjetos en cada uno de ellos.

Para el análisis funcional de la cNLS, las células HeLa y 293T se transfectaron con el plásmido pEYFP-C1 a modo de control negativo; un plásmido con una NLS muy activa (la NLS del antígeno T del virus

SV40) a modo de control positivo y el plásmido con la cNLS en estudio (pEYFP-cNLS). En el caso de las cNESs, las células HeLa se transfectaron con el plásmido Rev 1.4-GFP como control negativo; un plásmido con una NES muy activa como control positivo (García-Santisteban *et al.*, 2012) y el plásmido con las cNESs en estudio (Rev 1.4-cNES-GFP). Las muestras para el ensayo de las cNESs se prepararon por duplicado para su posterior tratamiento con CHX y Act. D.

La transfección se realizó empleando el reactivo X-tremeGENE 9 (Roche Diagnostics), que forma un complejo con el DNA permitiendo su transporte al interior celular. A los pocillos con las células ya adheridas al cubreobjetos se les añadió 750 μ l de DMEM completo, 1 μ l del DNA correspondiente en cada caso y 37,5 μ l de una mezcla de X-tremeGENE 9 y DMEM sin FBS ni P/S en proporción 3:100. Veinticuatro horas después de la transfección, cada una de las muestras para el estudio de las cNESs fue incubada durante 3 horas en DMEM que contenía 10 μ g/ml CHX y 5 μ g/ml Act.D o solo CHX. Tras la incubación, las células se fijaron con 3,7% formaldehído en PBS y los cubreobjetos se montaron sobre los portaobjetos empleando el medio de montaje Vectashield con DAPI (Vector Laboratories) para su observación al microscopio.

2.2.5. Microscopía de fluorescencia

Para el análisis de las muestras se utilizó un microscopio de fluorescencia Zeiss Axioskop de los Servicios Generales de Investigación (SGIker) de la UPV/EHU. Las imágenes fueron tomadas con una cámara Nikon DS-Qi1Mc controlada mediante el software NIS-Elements F y con un objetivo 40X. En el caso de las cNESs se realizó un recuento “a ciegas” (sin conocer la identidad de la muestra) de aproximadamente 200 células para cada muestra. La señal fluorescente de cada célula se clasificó como nuclear, citoplásmica o nucleocitoplásmica.

3. RESULTADOS

3.1. Predicción y selección de potenciales señales de transporte nucleocitoplásmico en la proteína KMT2D mediante análisis bioinformático

3.1.1. Selección de una potencial NLS en KMT2D

Cada uno de los tres programas ofreció varias potenciales NLSs. Con el fin de aumentar las probabilidades de que las NLSs predichas fueran funcionales, se decidió seleccionar para su estudio aquellas secuencias que hubiesen sido identificadas por más de un programa. Wregex y ELM coincidieron en la predicción de una única señal. De las señales predichas por PSORT, una se solapaba con el último aminoácido de la secuencia identificada por los otros dos programas. Por ello, se seleccionó una posible cNLS para la clonación cuya secuencia incluía las detectadas por los tres programas (Tabla 1).

Tabla 1. Resultados de las predicciones de las posibles NLSs coincidentes en los programas Wregex, ELM y PSORT y posible cNLS seleccionada para el estudio. Se indican los aminoácidos de inicio y final de las señales detectadas, así como su secuencia aminoacídica. El código de colores permite ver el solapamiento entre las distintas secuencias predichas por los tres programas en la cNLS en estudio.

Programa	Inicio	Final	Secuencia
Wregex	1684	1689	S K K R K R
ELM	1684	1689	S K K R K R
PSORT	1689	1705	R K P Y R P G I G G F M V R Q R K
cNLS	1680	1708	E T E E S K K R K R K P Y R P G I G G F M V R Q R K S H T

3.1.2. Selección de dos potenciales NESs en KMT2D

Para la selección de las posibles NESs se utilizó el mismo criterio empleado en el caso de las NLSs (Tabla 2). Wregex y ELM volvieron a coincidir en una señal, que se seleccionó para la clonación y se denominó cNES 1. Además, ELM predijo otra NES cuya secuencia se solapaba con dos señales identificadas por el programa LocNES. Esta región, a la que se denominó cNES 2, también se seleccionó para su estudio experimental. El programa NetNES no predijo ninguna señal que coincidiese con las detectadas por los otros programas.

Tabla 2. Resultados de las predicciones de las posibles NESs coincidentes en los programas Wregex, ELM y LocNES y posibles cNES seleccionadas para el estudio (cNES 1, cNES 2). Se indican los aminoácidos de inicio y final de las señales detectadas, así como su secuencia aminoacídica. El código de colores permite ver el solapamiento entre las distintas secuencias predichas por los tres programas en las cNESs en estudio. En negrita y subrayado se indican los aminoácidos hidrofóbicos que se ajustan a la secuencia NES consenso.

Programa	Inicio	Final	Secuencia
Wregex	3007	3022	E E <u>L</u> A H <u>L</u> G <u>L</u> G <u>V</u> D <u>V</u> A K G D
ELM	3007	3022	E E <u>L</u> A H <u>L</u> G <u>L</u> G <u>V</u> D <u>V</u> A K G D
	4554	4566	Q <u>L</u> K Q E <u>L</u> S L <u>L</u> P <u>L</u> T E
LocNES	4547	4561	A S E A L L K Q <u>L</u> K Q E <u>L</u> S L
	4550	4564	A L L K Q <u>L</u> K Q E <u>L</u> S L <u>L</u> P <u>L</u>
cNES 1	3006	3023	D E E <u>L</u> A H <u>L</u> G <u>L</u> G <u>V</u> D <u>V</u> A K G D D
cNES 2	4549	4567	E A L L K Q <u>L</u> K Q E <u>L</u> S L <u>L</u> P <u>L</u> T E P

3.1.3. Una mutación clínicamente relevante en KMT2D afecta a una potencial NES

Para investigar una posible relación entre el transporte alterado de la proteína KMT2D y su relevancia patológica se buscaron mutaciones clínicamente relevantes que pudiesen provocar la pérdida de actividad de señales de importación o exportación nuclear. No se encontraron mutaciones de interés registradas ni en la secuencia de la cNLS, ni en la de la cNES 1. Sin embargo, en Ensembl se localizó una mutación registrada en dbSNP (NCBI) que consistía en la eliminación de tres aminoácidos de la secuencia de la cNES 2 (p.Lys4555_Gln4557del) (Tabla 3). Se decidió clonar la cNES 2 mutada junto con la cNLS, cNES 1 y cNES 2.

Tabla 3. Comparación entre la secuencia normal (wild type) de la cNES 2 y la secuencia con la mutación registrada en dbSNP (mutada). En rojo se muestran los tres aminoácidos que no están presentes en la secuencia mutada. En negrita y subrayado se indican los aminoácidos hidrofóbicos que se ajustan a la secuencia NES consenso.

cNES 2	Secuencia
Wild type	E A L L K Q <u>L</u> <u>K</u> <u>Q</u> E <u>L</u> S L <u>L</u> P <u>L</u> T E P
Mutada (p.Lys4555_Gln4557del)	E A L L K Q E <u>L</u> S L <u>L</u> P <u>L</u> T E P

3.2. Resultados de la clonación

3.2.1. Transformación de las bacterias

Como se esperaba, crecieron numerosas colonias bacterianas en las placas positivas, es decir, en aquellas en las que se habían añadido los insertos. Por tanto, en dichas colonias se encontraban bacterias que habían incorporado el plásmido pEYFP-cNLS (Figura 6, a) y bacterias con el plásmido Rev 1.4-cNES-GFP con la cNES 1, cNES 2 o cNES 2 mutada (Figura 6, c). Las pocas colonias que crecieron en los controles negativos (Figura 6, b y d) probablemente se deban a que dichas bacterias incorporaron vectores que no se llegaron a cortar con las endonucleasas.

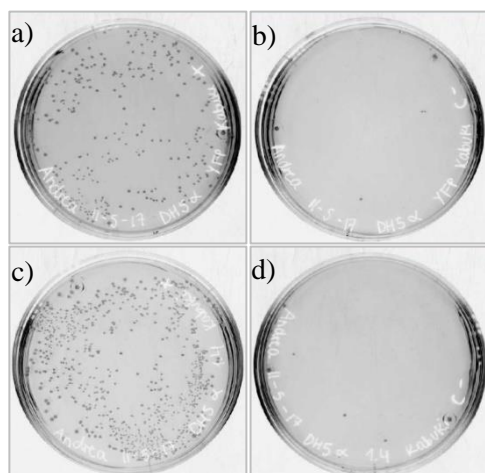


Figura 6. Resultado de la transformación de las bacterias *E.coli* DH5 α en placas con kanamicina. a) Bacterias que incorporaron el plásmido pEYFP-cNLS y b) control negativo. c) Bacterias que incorporaron el plásmido Rev 1.4-cNES-GFP con las posibles cNESs y d) control negativo.

3.2.2. Análisis mediante PCR y electroforesis de los insertos presentes en los plásmidos

En el caso de pEYFP-cNLS, haciendo una comparación con el control negativo (plásmido pEYFP-C1) se pudo observar que los plásmidos de dos de las colonias seleccionadas (C1 y C3) incorporaron más de un inserto. En cambio, los plásmidos de las otras tres colonias (C2, C4 y C5) contenían un solo inserto, que en teoría se trata de la cNLS. En el caso de los plásmidos Rev 1.4-cNES-GFP, se empleó el plásmido Rev 1.4-GFP a modo de control negativo y el mismo plásmido con una NES como control positivo. Las bandas de las cinco colonias seleccionadas (C1, C2, C3, C4 y C5) se encontraron a la misma altura que la del control positivo, por lo que incorporaron un solo inserto (Figura 7). En este caso hay tres señales distintas que pueden haberse insertado, cNES 1, cNES 2 o cNES 2 mutada.

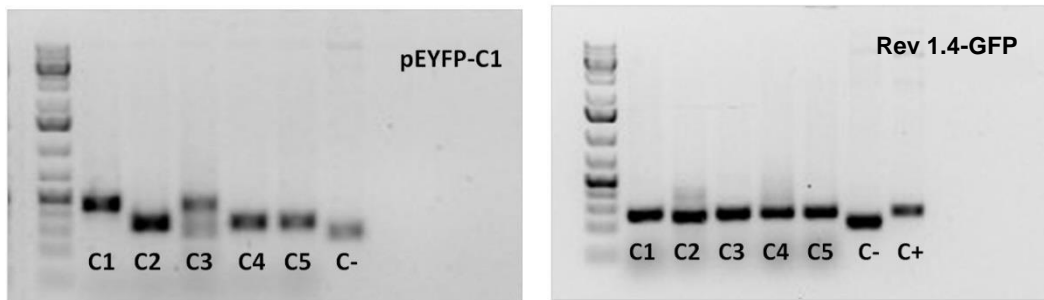


Figura 7. Resultados de la PCR y electroforesis en gel de agarosa 1,5% de los plásmidos pEYFP-cNLS y Rev 1.4-cNES-GFP extraídos de las colonias transformadas de E.coli DH5α. C1, C2, C3, C4 y C5 hacen referencia a las cinco muestras de colonias de las placas con kanamicina que se picaron en cada caso. El control negativo (C-) en ambos casos se trata del plásmido correspondiente sin inserto. En el caso del plásmido Rev 1.4-cNES-GFP se añadió también un control positivo (C+), que se trata del plásmido con una NES.

3.2.3. Secuenciación de los insertos

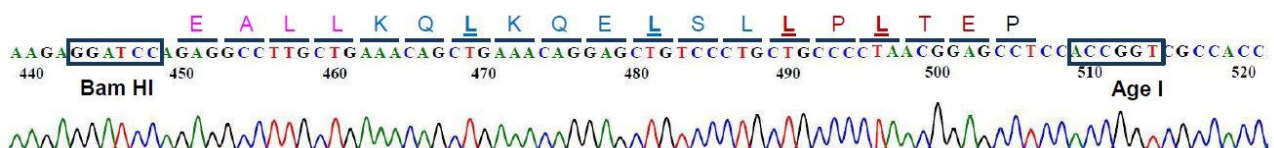
Los plásmidos de las cinco colonias seleccionadas con Rev 1.4-cNES-GFP y las colonias C4 y C5 con el plásmido pEYFP-cNLS fueron enviadas a secuenciar para asegurar que la señal se hubiese incorporado correctamente y, en el caso del plásmido Rev 1.4-cNES-GFP, también para comprobar cuál de las posibles señales se había insertado (Figura 8). La secuenciación de las dos muestras del plásmido pEYFP-cNLS demuestra que en ambos casos la cNLS se clonó correctamente. En el caso de Rev 1.4-cNES-GFP, las muestras C3 y C5 incorporaron la cNES 1 y la muestra C1 la cNES 2. Los plásmidos de C2 y C4 presentaron problemas en la secuenciación, probablemente debido a la mezcla de dos colonias al picar las bacterias. La cNES 2 mutada no se clonó correctamente en ninguna de las colonias seleccionadas, por lo que el estudio se continuó con el resto de señales (cNLS, cNES 1 y cNES 2).

a)

Plásmido	Muestra	Resultado
pEYFP-cNLS	C4	cNLS
	C5	cNLS
Rev 1.4-cNES-GFP	C1	cNES 2
	C2	-
	C3	cNES 1
	C4	-
	C5	cNES 1

Figura 8. Resultados obtenidos tras la secuenciación de los plásmidos. a) Se indican las señales insertadas en los diferentes plásmidos de las colonias seleccionadas. El signo (-) indica que no se pudo confirmar la secuencia insertada. b) Resultado de la secuenciación de la muestra C1 del plásmido Rev 1.4-cNES-GFP. Se pueden observar las dos secuencias de corte específicas para las endonucleasas Bam HI y Age I, así como la secuencia tanto nucleotídica como aminoacídica de la cNES 2. En negrita y subrayado se indican los aminoácidos hidrofóbicos que se ajustan a la secuencia NES consenso.

b)



3.3. La región 1680-1708 de la proteína KMT2D es una señal de localización nuclear funcional

Tras la transfección de las células, tanto HeLa como 293T, con los plásmidos correspondientes, se estudió la distribución subcelular de la proteína YFP mediante microscopía de fluorescencia. El control negativo

(plásmido pEYFP-C1) mostró fluorescencia tanto en el citoplasma como en el núcleo. En el control positivo solamente se observó fluorescencia en el núcleo, debido a la actividad importadora de la NLS fuerte fusionada a YFP. Las muestras de las cNLSs en estudio se compararon con las de los controles. Las células transfectadas con los plásmidos con la cNLS presentaron una clara acumulación de fluorescencia en el núcleo (Figura 9). Esto indica que la región 1680-1708 de KMT2D es una NLS funcional, ya que es capaz de dirigir la proteína YFP desde el citoplasma, donde había sido sintetizada, hasta el núcleo.

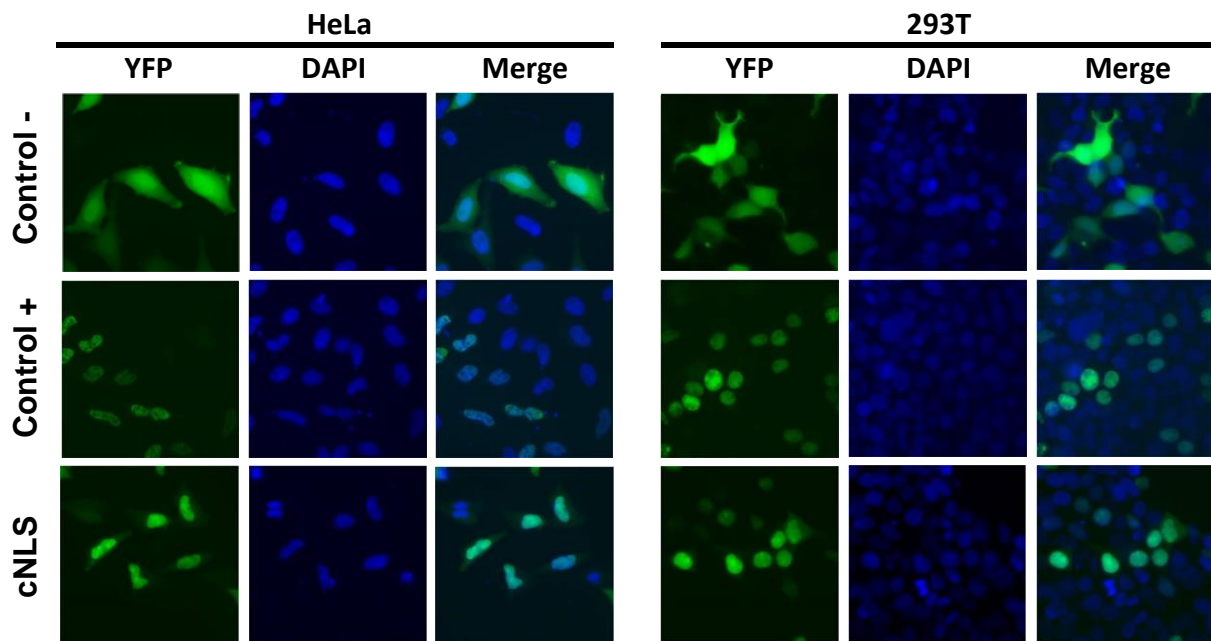


Figura 9. Resultados del análisis experimental de potenciales NLSs. Las imágenes obtenidas por microscopía de fluorescencia muestran la localización subcelular de la proteína YFP tras la transfección de los plásmidos pEYFP-cNLS en células HeLa y 293T. En azul se muestran los núcleos teñidos con DAPI. En ambas líneas celulares se observa que la cNLS dirige las proteínas al núcleo, por lo que se trata de una NLS funcional. Las imágenes fueron tomadas con un objetivo 40x.

3.4. La región 4549-4567 de la proteína KMT2D funciona como una señal de exportación nuclear débil

Una vez las células HeLa fueron transfectadas con los distintos plásmidos y tratadas con CHX o CHX y Act. D, se estudió la distribución subcelular de la proteína Rev 1.4-cNES-GFP en cada uno de los casos. En la figura 10 se observan las imágenes obtenidas por microscopía de fluorescencia de las distintas muestras en ausencia y presencia de Act. D, así como sus correspondientes gráficas, en las que se muestra el porcentaje de células que presentaron fluorescencia en el núcleo, en el citoplasma o en ambos. En todos los casos se pudo observar que al añadir Act. D el porcentaje de células que presentaron fluorescencia en el citoplasma aumentó, debido a su acción como inhibidor de la importación. El control negativo, que contenía el plásmido vacío, presentó fluorescencia mayoritariamente en el núcleo. Sin embargo, se observaron células que también presentaron fluorescencia en el citoplasma. De hecho, tras la incubación con Act. D el porcentaje de células que presentaron fluorescencia nucleocitoplásmica fue muy alto para

tratarse de un control negativo. En el caso del control positivo, que contenía una NES fuerte, tal y como se esperaba, se observó que prácticamente todas las células presentaron fluorescencia en el citoplasma.

En cuanto a las señales en estudio, los resultados obtenidos con la cNES 1 fueron prácticamente idénticos a los del control negativo, por lo que se concluyó que la cNES 1 no es una señal funcional. Por otro lado, en el caso de la cNES 2 en ausencia de Act. D, aunque muchas células presentaron fluorescencia solo en el núcleo, se observó un porcentaje notable de células que presentaron fluorescencia tanto en el núcleo como en el citoplasma respecto a las muestras de la cNES 1 y el control negativo. Esta diferencia fue mayor tras el tratamiento de las células con Act. D, llegando incluso a observarse un pequeño porcentaje de células que presentaron fluorescencia solo en el citoplasma. No obstante, la fluorescencia citoplásmica inducida por la cNES 2 fue muy inferior a la inducida por la NES fuerte del control positivo. Por todo ello, concluimos que la cNES 2, que abarca la región 4549-4567 de KMT2D, funciona como una señal de exportación nuclear débil.

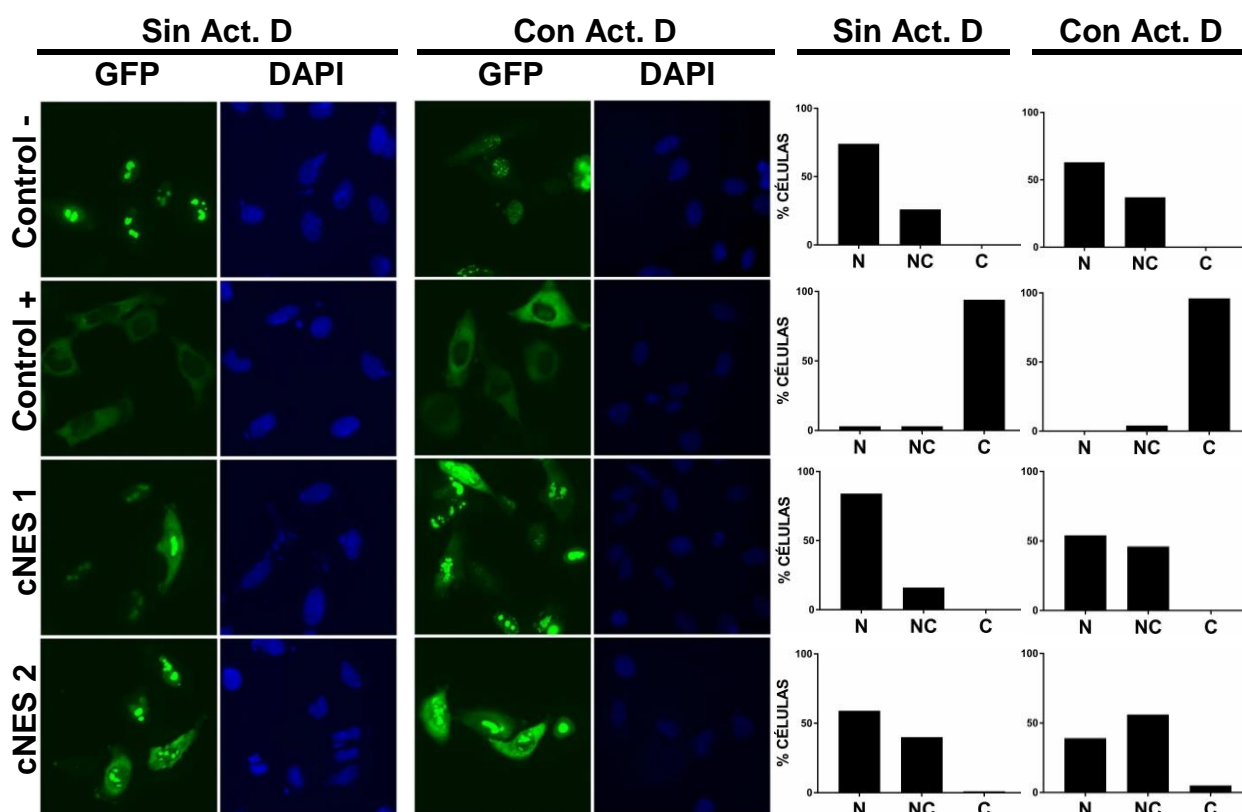


Figura 10. Resultados del ensayo de exportación nuclear de las NESs en estudio. En las imágenes obtenidas por microscopía de fluorescencia se muestra la localización subcelular de la proteína Rev 1.4-cNES-GFP tras la transfección de los plásmidos en células HeLa, tanto en ausencia de Act. D como tras la incubación durante 3 horas con Act. D. En azul se observan los núcleos teñidos con DAPI. Las imágenes fueron tomadas con un objetivo 40x. En las gráficas se representa el porcentaje de células que presentaron fluorescencia en el núcleo (N), en el citoplasma (C) o en ambos (NC) en cada uno de los casos. La actividad exportadora de la cNES 1 es similar a la observada en el control negativo, por lo que se considera que no es una verdadera NES. La cNES 2 muestra actividad, pero muy inferior a la del control positivo, lo que indica que se trata de una NES funcional débil.

4. DISCUSIÓN

Mediante una combinación de análisis bioinformáticos y experimentales hemos identificado dos nuevas señales de transporte nucleocitoplásmico funcionales (una NLS y una NES) en la proteína KMT2D.

Tras analizar las predicciones de las potenciales NLSs se decidió clonar una única cNLS, que comprendía las secuencias predichas por los tres programas (Wregex, ELM, PSORT) y que abarcaba la región 1680-1708 de la proteína. Wregex y ELM predijeron la secuencia SKKRKR, que *a priori* parece una NLS clásica monopartita, pues se trata de un *cluster* de aminoácidos cargados positivamente. PSORT predijo una señal que se solapaba con la última arginina de la otra secuencia, siendo ésta RKPYPGIGGFVVRQRK. Por tanto, la cNLS que se seleccionó para los ensayos podría ser una NLS clásica bipartita con dos *clusters* de argininas y lisinas separados por una región *linker*. El análisis experimental reveló que se trata de una NLS funcional, ya que las células transfectadas con los plásmidos con dicha señal presentaron abundante fluorescencia en el núcleo. Al tratarse de una NLS clásica, nuestros resultados indican que la proteína KMT2D es posiblemente transportada al núcleo mediante su interacción con la proteína adaptadora Imp- α y el receptor de transporte Imp- β 1 (Cautain *et al.*, 2015). Sin embargo, no se puede descartar que esta proteína también siga otros mecanismos de transporte alternativos.

En cuanto a las predicciones de las potenciales NESs, los programas Wregex y ELM coincidieron en una señal a la que se denominó cNES 1 (DEELAHGLGVDAKGGDD) y que estaba comprendida en la región 3006-3023 de la proteína. ELM y LocNES coincidieron en otra región a la que se llamó cNES 2 (EALLKQLKQELSLPLTEP; 4549-4567). Ambas cNESs se ajustan a la secuencia NES consenso caracterizada por una serie de residuos hidrofóbicos conservados (en este caso L y V) separados por un número variable de aminoácidos (Prieto *et al.*, 2014). Además de estas dos señales, también se seleccionó para el estudio la cNES 2 con una mutación registrada en dbSNP (NCBI). Esta mutación consiste en una delección de tres aminoácidos que incluye uno de los residuos hidrofóbicos conservados. El objetivo de incluir esta cNES 2 mutada en el ensayo fue comprobar si, en caso de ser la cNES 2 funcional, esta mutación impedía que se diese la exportación de la proteína. De este modo, se podría establecer una relación entre una posible patología y el transporte alterado de KMT2D como consecuencia de dicha mutación. Sin embargo, tras analizar los resultados de la secuenciación de los insertos de los plásmidos se observó que no se había conseguido clonar la cNES 2 mutada, por lo que se decidió continuar el estudio con las otras señales. Los resultados obtenidos en los ensayos de exportación nuclear (Henderson y Eleftheriou, 2000) indican que la cNES 1 no es una señal funcional, ya que incluso en presencia de Act. D la localización subcelular de la proteína Rev 1.4-cNES-GFP era muy similar a la del control negativo sin NES. Al comparar los resultados de la cNES2 con el control positivo era evidente que no se trata de una NES fuerte, pero la diferencia en el porcentaje de células con fluorescencia en el citoplasma o en el núcleo y el citoplasma respecto al control negativo sugiere que esta secuencia funciona como una NES débil. El hecho de que KMT2D no tenga una NES fuerte es bastante lógico teniendo en cuenta que desempeña su función en el núcleo celular. Por tanto, la presencia de una NES débil sería suficiente para regular la

actividad de la proteína. Además, en condiciones fisiológicas la actividad exportadora de esta NES débil podría verse enmascarada por la actividad importadora de la NLS detectada, al ser ésta más fuerte. Esto lleva a pensar que KMT2D puede ser una proteína que una vez sintetizada es dirigida al núcleo y que no tiene necesidad de ser exportada al citoplasma.

En definitiva, consideramos que la señal de mayor relevancia detectada en este estudio es la NLS, que va a ser crucial para la interacción de KMT2D con las proteínas transportadoras y posterior translocación al núcleo. Es posible que, además, gracias a esta señal de importación, KMT2D sea capaz de transportar al núcleo otras proteínas del complejo que no tengan NLS.

5. CONCLUSIÓN Y FUTUROS ESTUDIOS

Con este trabajo se ha contribuido a la caracterización de una proteína de gran relevancia fisiológica y con implicaciones en distintas patologías mediante la identificación de una señal de importación nuclear (NLS) y de una señal de exportación débil (NES) en su secuencia. Como futuros estudios, se propone extender el análisis a la proteína entera y no solo estudiar las regiones en las que se encuentran las señales, para evaluar la importancia de las señales en el contexto fisiológico. Sin embargo, hay que tener en cuenta la dificultad que presentaría trabajar con la proteína entera debido a su gran tamaño (5537 aminoácidos). Además, sería interesante realizar el estudio de la NES 2 con la mutación registrada para analizar el efecto que pudiese tener en el transporte de la proteína.

Finalmente, si se identifican nuevas mutaciones relacionadas con patologías en las regiones en las que se encuentran estas señales en la secuencia de KMT2D, se podría investigar la relación del transporte alterado de la proteína con las diferentes patologías.

6. BIBLIOGRAFÍA

Cautain, B., Hill, R., de Pedro, N., Link, W. (2015). Components and regulation of nuclear transport processes. *The FEBS Journal*, 282, 445-462.

Cheon, C. K., Ko, J. M. (2015). Kabuki syndrome: clinical and molecular characteristics. *Korean Journal of Pediatrics*, 58, 317-324.

Digilio, M. C. *et al.* (2017). Congenital heart defects in molecularly proven Kabuki syndrome patients. *American Journal of Medical Genetics*, 173, 2912-2922.

Dinkel, H. *et al.* (2016). ELM 2016—data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Research*, 44, 294-300.

Froimchuk, E., Jang, Y., Ge, K. (2017). Histone H3 lysine 4 methyltransferase KMT2D. *Gene*, 627, 337-342.

García-Santisteban, I., Bañuelos, S., Rodríguez, J. A. (2012). A global survey of CRM1-dependent nuclear export sequences in the human deubiquitinase family. *Biochemical Journal*, 441, 209-217.

Guo, C. *et al.* (2013). KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. *Oncotarget*, 4, 2144-2153.

Güttler, T. *et al.* (2010). NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nature Structural & Molecular Biology*, 17, 1367-1377.

- Henderson, B.R., Eleftheriou, A. (2000). A comparison of the activity, sequence specificity and CRM1-dependence of different nuclear export signals. *Experimental Cell Research*, 256, 213-224.
- la Cour, T., Kiemer, L., Mølgaard, A., Gupta, R., Skriver, K., Brunak, S. (2004). Analysis and prediction of leucine-rich nuclear export signals. *Protein Engineering, Design and Selection*, 17, 527-536.
- Nakai, K., Horton, P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends in Biochemical Sciences*, 24, 34-36.
- Prieto, G., Fullaondo, A., Rodriguez, J. A. (2014). Prediction of nuclear export signals using weighted regular expressions (Wregex). *Bioinformatics*, 30, 1220-1227.
- Rao, R. C., Dou, Y. (2015). Hijacked in cancer: the MLL/KMT2 family of methyltransferases. *Nature Reviews Cancer*, 15, 334-346.
- Teranishi, H. *et al.* (2018). Cancer management in Kabuki Syndrome: the first case of Wilms tumor and a literature review. *Journal of Pediatric Hematology/ Oncology*, [Epub ahead of print].
- Toska, E. *et al.* (2017). PI3K pathway regulates ER-dependent transcription in breast cancer through the epigenetic regulator KMT2D. *Science*, 355, 1324-1330.
- Xu, D. *et al.* (2015). LocNES: a computational tool for locating classical NESs in CRM1 cargo proteins. *Bioinformatics*, 31, 1357-1365.