

# Diseño y predicción de promotores específicos de neuroblastoma.

Jorge Mínguez Martínez<sup>1, 2, a</sup>, Raúl Torres Ruíz<sup>2</sup>, Sandra Rodríguez Perales<sup>2, b</sup>

1. Unidad de Bioquímica y Biología Molecular, Departamento de Biología de Sistemas, Facultad de Medicina y Ciencias de la Salud, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, España. 2. Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid, España.

a. [jorge\\_minguez@hotmail.com](mailto:jorge_minguez@hotmail.com) b. [srodriguezp@cnio.es](mailto:srodriguezp@cnio.es)

**Palabras clave:** neuroblastoma; promotores específicos; RT-qPCR; clonaje de promotores; citometría

## Resumen

El neuroblastoma es una enfermedad infantil caracterizada por el desarrollo de tumores a partir de células cancerígenas del tejido nervioso de las glándulas suprarrenales. Existen marcadores genéticos para caracterizar, clasificar y diagnosticar este tipo de tumores. Un ejemplo de estos marcadores es la amplificación de MYCN, frecuentemente utilizada como herramienta de pronóstico. A pesar de los tratamientos disponibles actualmente, el neuroblastoma es un tipo de cáncer que sigue estando caracterizado por una alta mortalidad. Por este motivo, es necesario el desarrollo de nuevas terapias que dirijan su acción específicamente a este tipo de células tumorales, maximizando su actividad en el tejido diana y reduciendo su toxicidad en otros tejidos. Un ejemplo de este tipo de terapia podría ser la edición génica. Para lograr una expresión específica en un tejido determinado, la expresión de genes terapéuticos debe estar controlada por promotores específicos de ese tejido. Para ello se deben analizar las secuencias promotoras que regulan la transcripción de genes cuya expresión sea restringida a neuroblastoma. Sin embargo, estas secuencias promotoras abarcan grandes regiones de ADN, y los vectores utilizados en terapia génica solamente pueden incorporar secuencias de un tamaño limitado. Por esta razón, se deben seleccionar los fragmentos esenciales que garantizan el funcionamiento del promotor. En este trabajo, se realiza un análisis y predicción de las secuencias fundamentales de cuatro promotores de proteínas expresadas diferencialmente en neuroblastoma para, posteriormente, evaluar su actividad mediante la medida de la expresión de una proteína reportera.

**Cita:** Mínguez Martínez, Jorge; Torres Ruíz, Raúl; Rodríguez Perales, Sandra (2019) Diseño y predicción de promotores específicos de neuroblastoma. *dianas* 8 (2): e201909fa05. ISSN 1886-8746 (electronic) [journal.dianas.e201909fa05](http://journal.dianas.e201909fa05) <http://www3.uah.es/dianas?e201909fa05>. URI <http://hdl.handle.net/10017/15181>

**Copyright:** © Mínguez-Martínez J, Torres-Ruiz R, Rodríguez-Perales S. Algunos derechos reservados. Este es un artículo open-access distribuido bajo los términos de una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional. <http://creativecommons.org/licenses/by-nc-nd/4.0/>

## Introducción

El neuroblastoma (NB) es un tipo de tumor sólido neuroendocrino extracranial que se desarrolla fundamentalmente en la infancia, donde en torno a un 90% de los casos se diagnostican antes de los diez años de edad [1]. Su desarrollo se produce a partir de las células precursoras del linaje simpatoadrenal, derivadas a su vez de la cresta neural [2]. Ésta es una estructura formada por un número limitado de células pluripotentes con capacidad migratoria que pueden dar lugar a distintos tipos de tejidos al recibir determinadas señales extracelulares [3].

Los pacientes de NB cursan con perfiles clínicos muy heterogéneos. Esto depende de muchos factores como, por ejemplo, la edad en la que se diagnostica, los marcadores genéticos, el lugar en el que se desarrolla el tumor primario o el grado de progresión, dado por el *International Neuroblastoma Pathological Classification* (INPC) y el *International Neuroblastoma Staging System* (INSS) [4, 5]. Los tumores primarios pueden producirse en cualquier parte del sistema nervioso simpático, pero es más frecuente encontrarlos en la médula adrenal y los ganglios paraespinales. Para caracterizar y diagnosticar a estos pacientes de NB desde el punto de vista genético, se han identificado una gran cantidad de marcadores que pueden ser utilizados adicionalmente como herramientas de pronóstico. Entre ellas podemos encontrar anomalías como pérdidas o ganancias de brazos cromosómicos – deleción 1p y 11q, ganancia 17q – amplificación de MYCN o una variación en la ploidía de las células [6-8]. Aproximadamente un 25% de los pacientes presentan amplificación del gen MYCN, íntimamente ligada a un pronóstico adverso. Actualmente, esta amplificación es el mejor marcador genético de riesgo de padecer NB.

El tratamiento de NB depende del riesgo de recaída del paciente. En los pacientes con pronóstico muy agresivo es frecuente utilizar una terapia multimodal que combine la eliminación quirúrgica del tumor, quimioterapia, trasplante autólogo de médula ósea, radioterapia, terapia de mantenimiento y, más recientemente, inmunoterapia. El tratamiento de estos pacientes continúa evolucionando a medida que se desarrollan nuevas opciones terapéuticas. Sin embargo, los pacientes clasificados como de alto riesgo (con presencia de metástasis y/o amplificación de MYCN), se asocian con una supervivencia del 29% a 5

años [9]. Asimismo, los pacientes que sufren recaída de NB raramente sobreviven. Por este motivo, este tipo de cáncer infantil sigue estando caracterizado por una alta mortalidad. La toxicidad y el alto potencial mutagénico de las terapias actuales han propiciado la búsqueda de nuevas terapias específicas, como el uso de virus oncolíticos contra las células tumorales o la inhibición de moléculas características de este tipo de tumor, como es la sobreexpresión de *MYCN* [10, 11].

Actualmente, se han intentado aplicar distintas tecnologías de edición genómica como posibles estrategias terapéuticas en muchas enfermedades humanas de origen genético [12]. Es el caso del sistema CRISPR-Cas9, el cual puede inducir eficientemente cortes de doble hebra en secuencias específicas del ADN [13, 14]. Varios estudios han demostrado que la aplicación de este sistema a células tumorales de NB es muy prometedora como herramienta terapéutica [15, 16]. Pero en la práctica, sin embargo, la principal dificultad se encuentra en dirigir estos sistemas de edición o limitar la expresión de sus componentes específicamente a las células tumorales.

La expresión de los genes eucarióticos depende fundamentalmente de una región promotora – la cual suele consistir en fragmentos relativamente grandes de ADN anteriores a la secuencia codificante – y de secuencias potenciadoras que pueden encontrarse muy alejadas en el genoma. Las secuencias incluidas en la región promotora son reconocidas por diferentes proteínas y factores, activadores y represores, que regulan la unión de la ARN polimerasa II y, por tanto, la expresión del gen. Por tanto, escoger el promotor adecuado, especialmente si se trata de un promotor específico de tejido, es un paso fundamental para conseguir la expresión dirigida de las distintas estrategias terapéuticas en las células diana. En general, estos promotores no solo proporcionan especificidad, sino que también disminuyen la toxicidad en otras localizaciones y maximiza la expresión en el tejido diana [17, 18].

Dado que la mayoría de vectores utilizados en terapia génica pueden incorporar *cassettes* de expresión de tamaño limitado, se debe realizar una búsqueda de la región esencial del promotor. Para conseguir la expresión activa de secuencias génicas con potencial terapéutico en células de NB, se deben evaluar promotores específicos que puedan ser caracterizados y seleccionar las regiones fundamentales para el funcionamiento del promotor. Tradicionalmente, esta selección se ha realizado mediante delecciones seriadas de la región anterior a la secuencia codificante. En la actualidad, se han llevado a cabo estudios de promotores de proteínas cuya expresión se limita a un subconjunto de tejidos derivados de la cresta neural, tejido de origen de NB, como NCX, THp o RAD51AP1 [19, 20]. Sin embargo, estos estudios han sido invalidados por otros que han demostrado falta de especificidad y pobres resultados. Como conclusión, no se han podido considerar aún promotores específicos de tejido que puedan ser utilizados en terapia génica contra NB [21, 22]. Con este fin, los objetivos principales que se abordarán en este trabajo son: (i) estudiar los perfiles de expresión de proteínas específicas de células tumorales de NB, (ii) analizar y predecir las regiones promotoras de los genes que codifican para estas proteínas mediante programas y plataformas bioinformáticas, y (iii) evaluar la fuerza de estas regiones promotoras mediante la medida de la expresión de un gen reportero.

## Materiales y métodos

### Búsqueda bibliográfica y selección de genes candidatos

La identificación y selección de genes candidatos se basó en trabajos publicados anteriormente cuyos resultados estuvieran registrados y confirmados en la base de datos *The Human Protein Atlas* ([www.proteinatlas.org/](http://www.proteinatlas.org/), consultado en abril de 2019). Adicionalmente, se recopilaron datos de estas proteínas y sus correspondientes genes a través de las plataformas *Uniprot* y *GenBank* ([www.uniprot.org/](http://www.uniprot.org/) y [www.ncbi.nlm.nih.gov/genbank/](http://www.ncbi.nlm.nih.gov/genbank/), consultado en abril de 2019).

### Líneas celulares utilizadas

Se utilizaron diez líneas celulares humanas: tres derivadas de NB (dos de ellas con amplificación de *MYCN*): SK-N-AS (ATCC CRL-2137, NB sin amplificación), IMR-32 (ATCC CCL-127, NB) y LAN-5 (DSMZ ACC673, NB); y el resto derivadas de distintos tipos de tumores y células primarias normales: A673 (ATCC CRL-1598, sarcoma de Ewing), CD34 (ATCC PCS-800-012, hematopoyéticas), O7N217 (glándula suprarrenal, cedidas por el Grupo de Cáncer Endocrino del Centro Nacional de Investigaciones Oncológicas), ASC52T (ATCC SCRC-4000, células mesenquimales), 2H (mesenquimales cedidas por el Clinic UB, Laboratori Institut Josep Carreras, Facultat de Medicina, Universidad de Barcelona), U-2 OS (ATCC HTB-96, osteosarcoma) y HEK293T (ATCC CRL-3216, embrionarias de riñón).

Para el cultivo de IMR-32 y LAN-5 se utilizó medio RPMI (Gibco) suplementado con aminoácidos no esenciales; para las células mesenquimales, medio Mesenpro RS (Gibco); y se utilizó medio *Dulbecco's Modified Eagle's Medium* (DMEM) (Sigma) para las demás líneas celulares. Todos ellos fueron suplementados con 10 % de suero bovino fetal (FBS), 1% de L-Glutamina 200mM (HyClone) y 1% de penicilina-estreptomicina 5000 U/mL (Gibco). Se cultivaron en un incubador Nuair 4950E, a 37°C, con concentraciones de 5% de CO<sub>2</sub> y 5% de O<sub>2</sub>. Los pases celulares se realizaron cuando las células

alcanzaban una confluencia del 80%, con lavado en *Phosphate Buffered Saline* (PBS) y levantando las células con 0,05% Trypsin-EDTA (Gibco).

## RT-qPCR

Las extracciones de ARN se realizaron mediante *TRIzol Reagent* (Invitrogen, Thermo Fisher) y el kit de extracción *RNeasy Mini* (QIAGEN). La retrotranscripción del ARN se realizó con la enzima *MultiScribe Reverse Transcriptase* (Invitrogen, Thermo Fisher) según las instrucciones del fabricante. Se comprobó el paso a ADN complementario mediante una reacción de PCR para la amplificación del gen *GAPDH* (30'' 98°C, 50'' 60°C, 50'' 72°C; 35 ciclos). El diseño de cebadores para qPCR se realizó a partir de las secuencias extraídas de Genome Browser (Tabla 1). Fueron analizados y comprobados mediante los programas web de *Primer-Blast* (NCBI) e *IDT RealTime PCR Tool* (IDT). Las reacciones de qPCR se realizaron por triplicado, en placa de 384 pocillos, utilizando *PowerSYBR Green PCR Master Mix* (Applied Biosystems, Thermo Fisher) y el gen *GUSB* como control interno de expresión. Su lectura se realizó en un dispositivo *QuantStudio 6 Flex Real-Time PCR System* (Applied Biosystems), y para el análisis de los resultados se utilizó el propio software del dispositivo.

Nombre	Secuencia	Nombre	Secuencia
Fw ASCL1	AGCTTCTCGACTTCACCAAC	Fw ARHGAP36	AAGACCGCCCATTGCTC
Rv ASCL1	CAACGCCACTGACAAGAAAG	Rv ARHGAP37	TGTTTCCTGTGACATCCCAG
Fw SCG3	CAACCAAGAATCGAAAACCTGATC	Fw PHOX2A	CCACTACCCCGACATTTACAC
Rv SCG3	CTGGCAGCGATTTTATGGAC	Rv PHOX2A	CTCCTGTTTGCGGAACTTG
Fw E2F1	CCGGGGAATGAAGGTGAACA	Fw NPY	CATCCGCTGGCTCTCACC
Rv E2F1	GAGCAAAAAGGGCCGAAAAGTG	Rv NPY	CCCAGTCGCTTGTTACCTAG
Fw ARF	GTTTTCGTGGTTCACATCCC	Fw SLC6A2	GTCCTGCCTTCTCTCTGTTC
Rv ARF	GCGCTGCCATCATCATG	Rv SLC6A2	CCTGCGTGCTGAGGAACTTA
Fw DBH	TGTCCACTTGGTCTACGGGA	Fw GUSB	GAAGAAGTGGTGCCTAGGG
Rv DBH	CTCCGGTTCGGGGATATTGG	Rv GUSB	AAGGATTTGGTGTGAGCGAT
Fw CHGB	ATGAAGGAATGGTGACTCGC	GAPDH Fwd	TCCAAAATCAAGTGGGGCGA
Rv CHGB	CAGTTGTCTCTTTGCTTTGACG	GAPDH Rev	TGATGACCCTTTGGCTCCC

Tabla 1.- Secuencias de los oligonucleótidos utilizados para la reacción de RT-qPCR de los genes seleccionados como candidatos.

## Predicción de secuencias promotoras

Se seleccionaron las secuencias que preceden a la secuencia codificante (ATG) de cada gen candidato desde *Genome Browser* (unas 10 kb) y se analizaron mediante los programas de *Segway Encyclopedia* y *Promoter 2.0* ([noble.gs.washington.edu/proj/encyclopedia](http://noble.gs.washington.edu/proj/encyclopedia) y [www.cbs.dtu.dk/services/Promoter/](http://www.cbs.dtu.dk/services/Promoter/), consultado en mayo de 2019).

## Diseño de oligonucleótidos, PCR y electroforesis

Para el diseño de los cebadores de las reacciones de PCR se utilizaron los programas *Primer-Blast* (NCBI) y las herramientas *UNAFold* y *OligoAnalyzer* (IDT) (Tabla 2). Una vez obtenidos los oligonucleótidos se resuspendieron a una concentración de 100µM, y fueron diluidos a 10µM para su uso. Las temperaturas de anillamiento (Tm) de cada cebador fueron calculadas mediante la herramienta *NEB Tm Calculator* (New England Biolabs).

Nombre	Secuencia
OligoEmpty_Fwd	AGCGCTGGATTTAATTAAGTTCTCGCTAGCGTCATAGTTCGAACTGACAATG
OligoEmpty_Rev	GATCCATTGTCAGTTCGAACTATGACGCTAGCGAGAAGTAAATTAATCCAGCGCT
E2F1_BamHI-Fwd	AATCGGATCCCCAAATCCTTTTGGCCGCGA
E2F1_PacI-Rev	GCATTTAATTAAGATTAGGATAATGTCATGGGTGAGG
PHOX2A_BamHI-Fwd	AATCGGATCCCTGAGCGCCCGAGAGTC
PHOX2A_PacI-Rev	GCATTTAATTAACCGTCCATGTCTCATCCCGA
CHGB_PacI-Fwd	GCATTTAATTAAGAAGATGCAGCTGTATATGTTCCA
CHGB_BamHI-Rev	AATCGGATCCTGTGCGGAAAGGAAGATGGC

Tabla 2.- Secuencias de los oligonucleótidos utilizados para las distintas reacciones de PCR (promotor sintético y promotores predichos de proteínas de NB) y para el inserto del plásmido pLV Empty Cherry RO (OligoEmpty).

Para la hibridación y fosforilación de los oligonucleótidos utilizados en la generación del plásmido pLV Empty Cherry RO (*Oligos Empty*), se utilizó el protocolo *Target Guide Sequence Cloning* (GeCKO, Zhang Lab).

Las reacciones de PCR de los promotores CHGB, E2F1 y PHOX2A se realizaron sobre ADN genómico humano. Se llevaron a cabo en un termociclador *Mastercycler epgradientS* (Eppendorf), con la enzima *Q5 DNA Polymerase*, el *Q5 Buffer Reaction* y el *Q5 Enhancer* (New England Biolabs; 30'' 98°C, 50'' 62°C, 50'' 72°C; 35 ciclos). Los productos se chequearon mediante gel de agarosa 1% con TBE1x, utilizando el marcador *MassRuler DNA Ladder Mix* (Thermo Fisher). La purificación de estos productos se realizó mediante el kit *QIAquick PCR Purification* (QIAGEN); y su cuantificación, en un dispositivo *Nanodrop ND- 1000* (ThermoScientific).

### Clonaje de plásmidos

Para las digestiones, tanto de vectores como de gBlocks y oligonucleótidos (IDT), se utilizaron las enzimas de restricción Afel, BamHI, BstBI, ClaI, KpnI, NheI y PacI, todas ellas con el buffer *CutSmart* (New England Biolabs), siguiendo las indicaciones del fabricante. Entre los vectores utilizados se incluyeron los plásmidos lentivirales pLV Tnt Cherry RO (plásmido disponible en el laboratorio que incluye la proteína fluorescente Cherry, resistencia a Puromicina y está orientado de manera inversa – *reverse oriented* –) y pLVX-Puro-mGFP-WDR34 (Addgene #64251, que contiene la secuencia *wild type* del promotor de citomegalovirus humano (CMV). El gBlock utilizado incluyó la secuencia promotora predicha de SCG3, y fue resuspendido a una concentración de 10ng/μL.

Dependiendo de la aplicación, la purificación de los productos obtenidos de las digestiones se realizó con el kit *QIAquick PCR Purification* (QIAGEN) para secuencias escindidas menores de 100 pb, y con el kit *QIAquick Gel Extraction* (QIAGEN) para secuencias mayores de 100 pb, previa electroforesis. La desfosforilación de estos productos se realizó mediante la enzima rSAP en buffer *CutSmart* (New England Biolabs). Las ligaciones se llevaron a cabo con la enzima *T4 DNA Ligase* y su correspondiente buffer (New England Biolabs), incubando durante 16 horas a 4°C. Los productos de ligación fueron transformados en bacterias *One Shot Stb13 Chemically Competent* (Invitrogen) por choque térmico a 42°C durante 45 segundos, seguido de recuperación en medio SOC (Invitrogen). Posteriormente, se sembraron en superficie y se seleccionaron en placas de medio Luria Broth (LB) con agar y ampicilina (Conda). De cada ligación, se llevó a cabo un control negativo de religado en el cual la ligación del vector desfosforilado se realizó sin inserto. Se eligieron varias colonias de cada transformación, en función del ratio con el religado, y se amplificaron en 3mL de LB líquido (Alaos) con 3μL de ampicilina (30μg/mL). Las extracciones de ADN plasmídico se realizaron con el kit *EZNA Plasmid DNA Mini* (OMEGA). Tras realizar una confirmación de la ligación por mapa de restricción (con enzimas para dianas dentro del inserto) en gel de agarosa 1%, se llevó a cabo una secuenciación Sanger (sistema ABI Prism, Applied Biosystems, Unidad de Genómica, Servicio de Secuenciación del CNIO). Los clones confirmados por secuenciación fueron amplificados en 200mL de LB líquido con 200μL de ampicilina y se realizaron maxi preparaciones de ADN plasmídico mediante el kit *Plasmid Maxi* (QIAGEN).

### Transfección con lipofectamina y citometría de flujo

La transfección se realizó sobre células IMR-32 y HEK293T al 70-80% de confluencia, mediante lipofectamina (Lipofectamine 2000, Invitrogen) siguiendo las instrucciones de la casa comercial. Tras la transfección con los distintos constructos, las células se incubaron durante 48 horas a 37°C. Posteriormente, se llevó a cabo un análisis por citometría de flujo en un dispositivo LSR FORTRESSA (Becton Dickinson), resuspendiendo las células en 200μL de PBS y añadiendo 1μL de DAPI. En este análisis se seleccionó la subpoblación celular que, excitada a una longitud de onda de 587nm, emitía a 610nm (expresión de la proteína roja fluorescente Cherry). Los resultados se analizaron mediante el programa *FlowJo*.

## Resultados

El objetivo principal de este estudio es la generación de herramientas (*cassettes* de expresión) para la expresión específica y restringida de genes terapéuticos en células tumorales de NB. Por tanto, se debe realizar: un estudio de la expresión de proteínas específicas de NB, un análisis y predicción de las regiones promotoras de los genes que codifican para estas proteínas, y una evaluación de su actividad y fuerza mediante la medida de la expresión del gen reportero *Cherry*.

### Estudio de expresión de proteínas específicas de NB

Se comenzó realizando una búsqueda exhaustiva de proteínas con expresión notable y diferencial en líneas celulares de NB. Esta búsqueda dirigida se realizó con distintas plataformas y bases de datos biológicos. Se centró en la plataforma *The Human Protein Atlas* contrastando con la bibliografía disponible. Como resultado, se identificaron un total de ocho proteínas candidatas con un nivel de expresión específico de líneas de NB: ASCL1, SCG3, DBH, CHGB, ARHGAP36, SLC6A2, PHOX2A y NPY. Además, se incluyeron dos proteínas con una expresión elevada en cáncer: E2F1 y ARF

Para comprobar su expresión diferencial, se realizaron reacciones de RT-qPCR en las que se testaron los diez genes candidatos, utilizando como control endógeno la expresión del gen *GUSB*, el cual codifica para la  $\beta$ -glucuronidasa (gen de expresión constitutiva). Se compararon los niveles de expresión en diez líneas celulares, tres de ellas de NB, con y sin amplificación de *MYCN* (SK-N-AS, IMR-32 y LAN-5), y un conjunto representativo de otros linajes celulares (CD34, O7N217, HEK293T) otros tipos de células tumorales (A673, U-2 OS) y algunas células primarias normales (ASC52T, 2H). Se diseñaron oligonucleótidos para llevar a cabo una reacción de PCR cuantitativa a partir de la secuencia de los genes que codifican para las proteínas seleccionadas. Como resultado de este análisis, se normalizaron los niveles de expresión respecto al control endógeno y se encontraron tres de los diez genes analizados cuya expresión era elevada y exclusiva de las líneas de NB: *SCG3*, *CHGB* y *PHOX2A*. Se encontró también una expresión elevada de *E2F1* para las líneas de NB y una expresión menor, aunque marcada, para las demás líneas tumorales. Además, se obtuvo una expresión relativa muy alta de varios genes (*ARHGAP36*, *SLC6A2*, *DBH* y *NPY*) en la línea O7N217 de glándula suprarrenal (Figura 1).

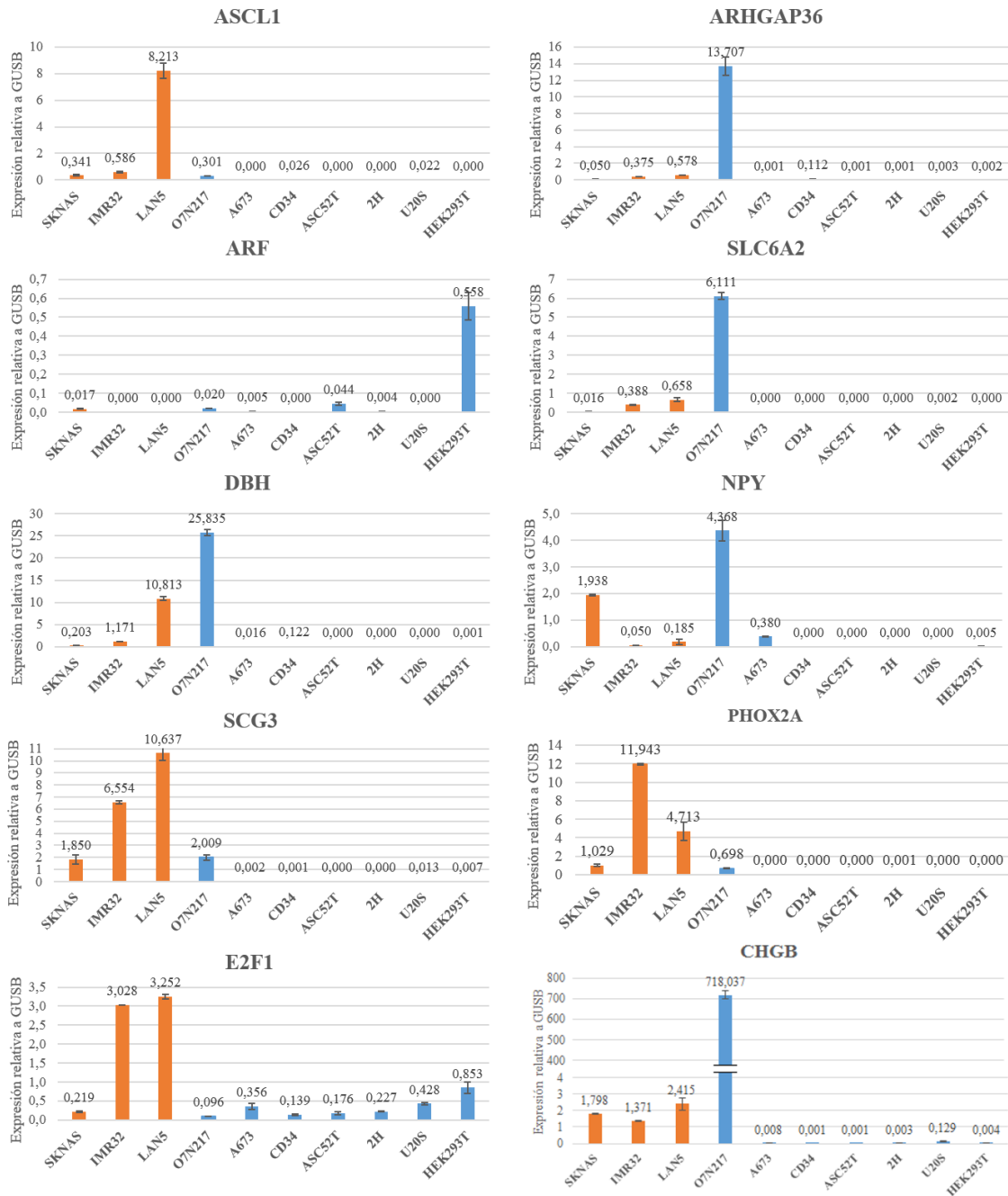


Figura 1.- Análisis de RT-qPCR. (Naranja: Líneas celulares de NB; Azul: Otras líneas celulares). Estudio de la expresión de diez proteínas diferentes (ASCL1, SCG3, E2F1, ARF, DBH, CHGB, ARHGAP36, SLC6A2, PHOX2A y NPY) en diez líneas celulares (SK-N-AS, IMR-32, LAN-5, O7N217, A673, CD34, ASC52T, 2H, U-2 OS y HEK293T). Las medidas obtenidas se normalizaron respecto a la expresión endógena de *GUSB*.

## Análisis y predicción de regiones promotoras

Una vez identificados los genes *SCG3*, *E2F1*, *CHGB* y *PHOX2A* con expresión potencialmente restringida a NB, se estudiaron las secuencias promotoras de los genes que codifican para las mismas. En primer lugar, se analizó la secuencia inmediatamente anterior al inicio de la región codificante (ATG) mediante el visualizador de *Genome Browser*, acotando además la búsqueda mediante la herramienta *ORegAnno*. Como resultado, se obtuvieron unas regiones amplias (alrededor de 10kb) que posteriormente se analizaron mediante dos programas de predicción de secuencias promotoras: *Segway Encyclopedia* y *Promoter 2.0.*, los cuales tienen en cuenta distintas características de promotores. Como resultado, y dado el límite de capacidad de los plásmidos de expresión, se seleccionaron secuencias de 950 pb para cada gen: *SCG3* (chr15: 51680805-51681755), *E2F1* (chr20: 33686266-33686716), *CHGB* (chr20: 5910683-5911633) y *PHOX2A* (chr11: 72244005-72244955).

## Medida de la actividad de las regiones promotoras. Clonaje, transfección y citometría.

Por una parte, una vez definidas las cuatro potenciales regiones promotoras, se optó por dos estrategias para su clonación: la síntesis de la secuencia seleccionada por una casa comercial (en formato gBlock; promotor de *SCG3*), o la síntesis de la región por PCR (*E2F1*, *PHOX2A* y *CHGB*) (Figura 2A). En ambos casos, tanto para gBlock como para los cebadores para la amplificación por PCR, se diseñaron incorporando dianas de restricción, para BamHI y PacI, en cada uno de los extremos para su posterior ligación con el vector de expresión. Por otra parte, como control positivo del estudio de expresión se utilizó el promotor CMV, ubicuo y potente, aislado a partir del plásmido pLVX-Puro-mGFP-WDR34 (Material suplementario C), mediante digestión con las enzimas BamHI y ClaI (esta última compatible con BstBI).

En este proyecto, se utilizó un vector de expresión para genes de mamíferos, el plásmido pLV Tnt Cherry RO (Material suplementario A). En este plásmido, la expresión de la proteína roja fluorescente Cherry está controlada por un promotor de partida Tnt. La estrategia seguida fue la sustitución de este promotor inicial por los distintos promotores obtenidos: los cuatro promotores a testar y el promotor ubicuo de CMV, de forma que la medida de la expresión de la proteína Cherry se pueda relacionar directamente con la potencia del promotor.

Se comenzó sustituyendo el promotor Tnt, mediante la digestión con BamHI y AfeI, por unos oligonucleótidos hibridados (*Oligos Empty*, Material Suplementario B) que dejaban extremos compatibles para el clonaje. Esta secuencia introducida se diseñó de forma que incorporara un *multicloning site*, con distintas dianas de restricción para las enzimas BamHI, BstBI, NheI, PacI y AfeI. Se obtuvo así un plásmido base al que se denominó pLV Empty Cherry RO en el que poder clonar las distintas secuencias promotoras. La digestión de pLV Empty Cherry RO se realizó, por una parte, con las enzimas BamHI y PacI para incorporar las secuencias promotoras predichas de *SCG3* (gBlock), *E2F1*, *CHGB* y *PHOX2A* (PCR); y, por otra parte, con BamHI y BstBI para incorporar la secuencia promotora de CMV. Una vez realizada la ligación y la transformación, se obtuvieron clones de *E. coli* (Stbl3) portadores de cada uno de los plásmidos resultantes, de los que se extrajo ADN. El correcto clonaje se confirmó mediante digestiones con NheI y KpnI (Figura 2B).

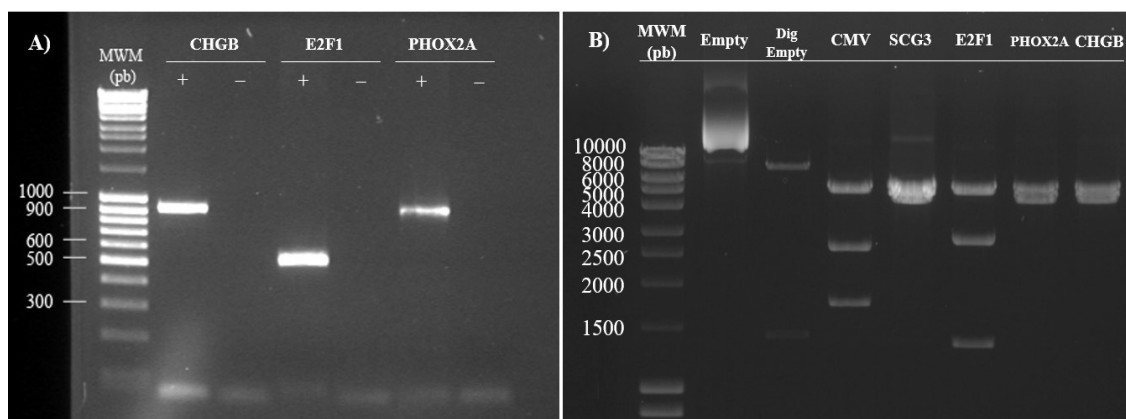


Figura 2.- A) Electroforesis de la reacción de PCR de las secuencias promotoras de *CHGB*, *E2F1* y *PHOX2A*. De izquierda a derecha: amplificación de los genes *CHGB* (906 pb), *E2F1* (500 pb) y *PHOX2A* (867 pb); cada uno con sus respectivos controles negativos. MWM: Molecular Weight Marker. B) Digestiones de plásmidos y clonajes realizados. De izquierda a derecha: pLV Empty Cherry RO (8931pb) y su digestión (diana interna) con NheI (1428 y 7503 pb). Digestiones con KpnI de los clonajes en pLV Empty Cherry RO de: *CMV wt* (1757, 2626 y 5218pb), *SCG3* (4631 y 5218pb), *E2F1* (1365, 2815 y 5218pb), *PHOX2A* (4547 y 5218pb) y *CHGB* (4725 y 5218pb).

A continuación, se transfectaron los diferentes constructos mediante lipofección en células IMR-32 (línea de NB con amplificación en *MYCN*) y en HEK293T, utilizando como control positivo de expresión el promotor de CMV y como control negativo el propio vector pLV Empty Cherry RO en el que no se incorporó ninguna secuencia promotora. Posteriormente, se realizó un análisis de la expresión de Cherry por citometría de flujo, en el que se identificó la expresión mínima como la del control negativo y la expresión máxima la del control positivo. Se obtuvo así una expresión aumentada en las células HEK293T para el promotor de *CHGB*, y en menor medida para los promotores de *E2F1* y *SCG3* (Figura 3).

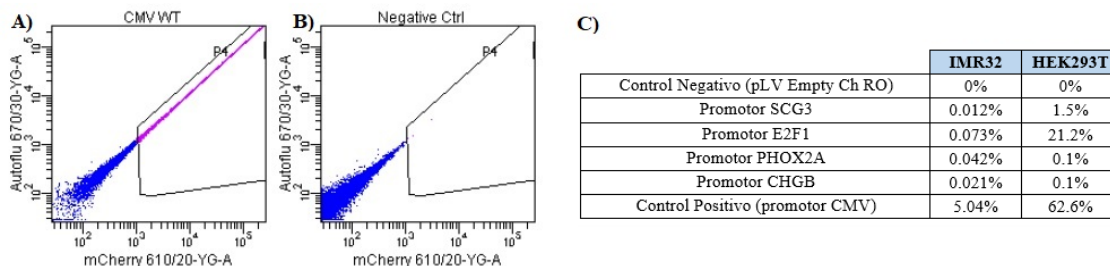


Figura 3.- A) Gráfica de citometría de flujo del control positivo de expresión, CMV. B) Gráfica de citometría de flujo del control negativo de expresión. C) Tabla resumen de los porcentajes de células vivas que presentan fluorescencia por la expresión de la proteína Cherry en función del promotor incorporado.

## Discusión

Las secuencias promotoras de los genes eucarióticos comprenden amplias regiones que regulan su expresión. Dentro de estas regiones, hay secuencias esenciales de reconocimiento de distintos factores activadores que aseguran el funcionamiento del promotor. Ya que los vectores utilizados pueden incorporar insertos de un tamaño determinado, la longitud final del fragmento a estudiar y evaluar es limitada. Por este motivo, la selección de las secuencias promotoras debe estar dirigida y coincidir con las regiones esenciales que garantizan la actividad del promotor. Para lograr restringir la expresión de genes terapéuticos a un tipo celular específico, debemos analizar los perfiles de expresión de distintas proteínas en dichas células, estudiar sus secuencias promotoras fundamentales y clonaras en vectores en los cuales se puedan evaluar sus niveles de expresión.

Las plataformas bioinformáticas que contienen información sobre la expresión de determinados genes, como *The Human Protein Atlas*, son de gran utilidad para la identificación de proteínas específicas de un determinado tejido o tipo celular, como en este caso células tumorales de NB. *The Human Protein Atlas* es una plataforma en la cual se puede encontrar información relativa a los perfiles de expresión de genes humanos, a nivel de ARNm y de proteína. Engloba datos de expresión analizados en un panel de 64 líneas celulares de distintos órganos y tejidos. Dado que una de estas líneas, SH-SY5Y, pertenece a NB, se seleccionaron diez proteínas candidatas cuya expresión se encontrara aumentada en la misma. También se estudiaron otras bases de datos en las que se reportaba información de expresión en tejidos primarios probable origen de NB.

El origen de este tipo de tumor son las células precursoras del linaje simpatoadrenal (derivadas a su vez de la cresta neural). Los ocho promotores seleccionados pertenecían a genes relacionados tanto con neuropéptidos y almacenamiento de neurotransmisores, como de precursores de péptidos regulatorios y proteínas relacionadas con la proliferación, por lo que se ajustaba al tipo de tejido. Por este motivo, tampoco fue de extrañar que la línea O7N217 (derivada de glándula suprarrenal) presentara una expresión aumentada para varias de las proteínas seleccionadas. No obstante, algunas de estas proteínas no fueron específicas de NB, sino que presentan un perfil de expresión modificado en cáncer. Por ejemplo, E2F1 es un factor para el control del ciclo celular por su unión a retinoblastoma, y ARF es una proteína que estabiliza a p53. En el estudio de expresión por PCR cuantitativa, la presencia de E2F1 en todas las líneas analizadas, aunque especialmente en NB, se observó como un marcador de crecimiento, debido a su papel en proliferación, replicación del ADN y diferenciación. Por otra parte, la expresión aumentada de los genes *SCG3*, *E2F1*, *PHOX2A* y *CHGB* en las tres líneas de NB testadas, con o sin amplificación, fue determinante para continuar con el estudio de sus secuencias promotoras.

Para el análisis y predicción de las regiones promotoras de las proteínas seleccionadas fue imprescindible el uso de herramientas computacionales, a través de las cuales se definieron claramente los criterios a seguir para la identificación de promotores. Las grandes regiones seleccionadas a partir de *Genome Browser* para su análisis tuvieron en cuenta características propias de los promotores. Al ser regiones del ADN menos condensadas (accesibles a la maquinaria de transcripción) son, por tanto, sitios hipersensibles a la digestión por DNasa I. También son buenos indicadores la presencia de islas CpG y la existencia de un gran número de secuencias consenso de reconocimiento por factores de transcripción. Por otra parte, la predicción de secuencias promotoras mediante las herramientas *Segway Encyclopedia* y

*Promoter 2.0.* fue decisiva, ya que acotó las secuencias seleccionadas. Mediante estos programas se buscó, a partir de la secuencia introducida, posibles sitios de interacción de distintos factores de transcripción con el ADN que pudieran estar involucrados en el reconocimiento de la región por la ARN polimerasa II y que, por tanto, pudieran formar parte de un promotor. Estos programas fueron capaces de distinguir regiones con mayor probabilidad de ser promotoras, regiones transcritas, zonas inactivas o silenciadas, heterocromatina e incluso elementos reguladores distales (*enhancers*).

Una vez obtenidas y clonadas las secuencias promotoras en un vector de expresión, se evaluó la fuerza de los distintos promotores. Esta evaluación debe estar caracterizada por un gen reportero cuya expresión pueda ser medida y así relacionarse proporcionalmente con la actividad de su región promotora. En este caso, se escogió la medida de la proteína roja fluorescente Cherry. Sus niveles de expresión en función del promotor testado se normalizaron respecto a la expresión ubicua controlada por el promotor de *CMV* como control positivo, y la expresión nula de pLV Empty Cherry RO (debido a la ausencia de promotor), como control negativo. A pesar de los perfiles de expresión dados para la línea IMR-32, los resultados obtenidos demuestran que las secuencias promotoras seleccionadas y clonadas presentan actividad promotora y son capaces de activar la expresión de la proteína Cherry en las células HEK293T. El promotor de *E2F1*, proteína característica de proliferación, crecimiento y cáncer, obtuvo los niveles más altos de expresión para la línea HEK293T, manteniéndose mínimas las obtenidas con los promotores específicos de NB. Dados los niveles de expresión del promotor de *CMV* en la línea IMR-32, sería necesario poner a punto el método de transfección de los distintos constructos, e implementar técnicas que aseguren una expresión no transitoria.

A pesar de esta evaluación, este proyecto se considera un estudio muy preliminar de promotores específicos para NB, por lo que los resultados obtenidos en este trabajo no son concluyentes. En estudios posteriores, estos análisis deberían incluir varias réplicas y ser realizados para todas las líneas de NB con las que se trabaja en el laboratorio. Unos resultados positivos implicarían la expresión diferencial de Cherry bajo el control de los promotores estudiados, comparando en todo momento respecto a la línea celular control. En cambio, unos resultados negativos podrían deberse a errores en la predicción y obtención de las secuencias promotoras.

### Próximos objetivos

Para futuras aproximaciones terapéuticas, se debe incluir la producción, mediante plásmidos estructurales, de vectores no replicativos lentivirales (LV) o adenoasociados (AAV), que encapsiden los distintos constructos. Estos constructos incorporarían, entre otros componentes, genes terapéuticos cuya expresión estuviera regulada por las regiones promotoras analizadas en este trabajo, de forma que se lograra una expresión específica cuando se tratara de células tumorales de NB. Esta transducción, realizada a un ratio (MOI) de infección bajo, garantizaría un control del número de copias de plásmido por célula – al contrario que la lipofección – y una expresión mantenida en el tiempo.

Por último, también se deberían estudiar otro tipo de estrategias más novedosas. Una aproximación interesante sería la elaboración de un promotor sintético que combine distintas secuencias de unión a factores de transcripción específicos o sobreexpresados en células de NB, hibridados en distintas combinaciones y con distinto grado de espaciamiento entre ellos. Como producto de esas hibridaciones, se generaría un amplio rango de tamaños que podrían ser amplificados mediante PCR para su posterior selección, aislamiento y clonaje.

## Conclusiones

1. Se han conseguido identificar cuatro proteínas (SCG3, E2F1, PHOX2A y CHGB) cuya expresión es específica dentro de las líneas celulares analizadas.
2. Se han logrado predecir, diseñar y obtener, mediante distintas estrategias las secuencias promotoras esenciales que regulan la expresión de genes que codifican para estas proteínas.
3. El estudio de la actividad promotora mediante citometría de flujo indica que las secuencias clonadas son capaces de activar la expresión de la proteína Cherry
4. Sin embargo, este estudio constituye un análisis muy preliminar que debería ser extendido en profundidad para conseguir datos de mayor poder estadístico.

## Bibliografía

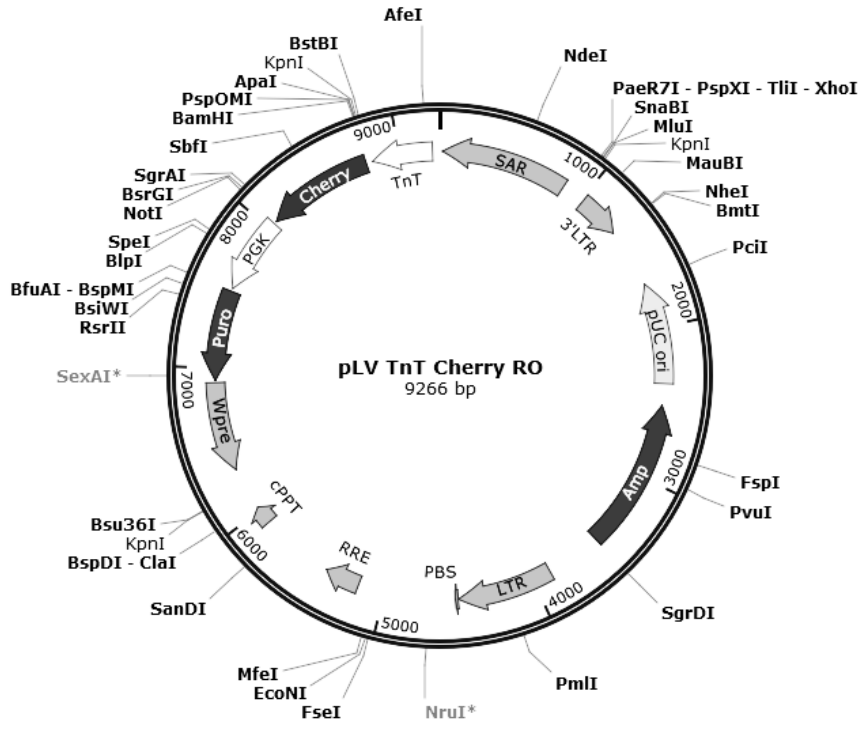
1. London, W. B., Castleberry, R. P., Matthay, et al. (2005). *Evidence for Age Cutoff Greater Than 365 Days for Neuroblastoma Risk Group Stratification in the Children's Oncology Group*. J Clin Oncol, 23(27), 6459–6465.
2. Matthay, K. K., Maris, J. M., Schleiermacher, G., et al. (2016). *Neuroblastoma*. Nat Rev Dis Primers, 2, 16078.



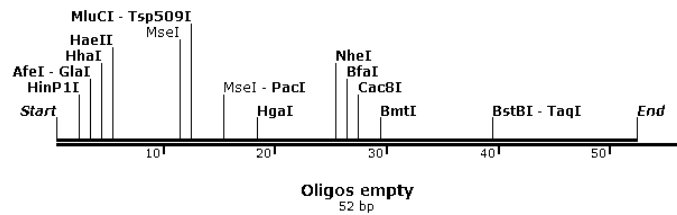
3. Mayor, R., & Theveneau, E. (2013). *The neural crest*. Development, 140(11), 2247–2251.
4. Shimada, H., Ambros, I. M., Dehner, et al. (1999). *The International Neuroblastoma Pathology Classification (the Shimada system)*. Cancer, 86(2), 364–372.
5. M., Pritchard, J., Berthold, F., Carlsen, et al. (1993). *Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment*. J Clin Oncol, 11(8), 1466–1477.
6. I, S., Łastowska, M., O’Neill, S., Pearson, A. D. J., et al. (1999). *Gain of Chromosome Arm 17q and Adverse Outcome in Patients with Neuroblastoma*. N. Engl. J. Med., 340(25), 1954–1961.
7. London, W. B., Mossé, Y. P., Wang, Q., et al. (2005). *Chromosome 1p and 11q Deletions and Outcome in Neuroblastoma*. N. Engl. J. Med., 353(21), 2243–2253.
8. Campbell, K., Shyr, D., Bagatell, R., et al. (2019). *Comprehensive evaluation of context dependence of the prognostic impact of MYCN amplification in neuroblastoma: A report from the International Neuroblastoma Risk Group (INRG) project*. Pediatric Blood & Cancer, e27819.
9. Castel, V., Cañete, A., Navarro, S. et al. (2001). *Outcome of high-risk neuroblastoma using a dose intensity approach: Improvement in initial but not in long-term results*. Med Pediatr Oncol., 37(6), 537–542.
10. Ramírez, M., García-Castro, J., Alemany, R. *Oncolytic Virotherapy for Neuroblastoma*. (2010) Discovery Medicine, 10(54), 387-93.
11. Kang, J.-H., Rychahou, P. G., Ishola, T. A., et al. (2006). *MYCN silencing induces differentiation and apoptosis in human neuroblastoma cells*. Biochem Biophys Res Commun, 351(1), 192–197.
12. Aguirre, A. J., Meyers, R. M., Weir, B. A., et al. (2016). *Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting*. Cancer Discovery, 6(8), 914–929.
13. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., et al. (2005). *Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements*. J Mol Evol, 60(2), 174–182.
14. Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). *Genome engineering using the CRISPR-Cas9 system*. Nature Protocols, 8(11), 2281–2308.
15. Kiessling, M. K., Schuierer, S., Stertz, S., et al (2016). *Identification of oncogenic driver mutations by genome-wide CRISPR-Cas9 dropout screening*. BMC Genomics, 17(1).
16. Yoda, H., Inoue, T., Shinozaki, Y., et al. (2018). *Direct Targeting of MYCN Gene Amplification by Site-Specific DNA Alkylation in Neuroblastoma*. Cancer Research, 79(4), 830-840.
17. Zheng, C., & Baum, B. J. (2008). *Evaluation of Promoters for Use in Tissue-Specific Gene Delivery*. Gene Therapy Protocols, 205–219.
18. Chen, X., Scapa, J. E., Liu, D. X., et al. (2016). *Cancer-specific promoters for expression-targeted gene therapy: ran, brms1 and mcm5*. The Journal of Gene Medicine, 18(7), 89–101.
19. Narita M, Bahar R, Hatano M, et al. (2001) *Tissue-specific expression of a suicide gene for selective killing of neuroblastoma cells using a promoter region of the NCX gene*. Cancer Gene Ther. 8(12):997–1002.
20. Valentijn, L. J., Koster, J., Haneveld, F., et al (2012). *Functional MYCN signature predicts outcome of neuroblastoma irrespective of MYCN amplification*. Proc. Natl. Acad. Sci., 109(47), 19190–19195.
21. Steffens, S., Frank, S., Sandquist, A., et al (2002). *Selective Herpes Simplex Virus Thymidine Kinase Gene Expression Regulated by a Neuroblastoma-Specific Promoter*. Mol Ther, 5(5), S33–S34.
22. Tivnan, A. and Stallings, R.L. (2010) *Neuroblastoma-specific expression of potential therapeutics cannot be achieved using a promoter region of the NCX (TLX2) gene*. Cancer Gene Ther. 17(5): 373–374.

Material suplementario

A)



B)



w\_authors

w\_address

C)

