

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 862 526**

51 Int. Cl.:

C12N 15/864 (2006.01)

A61K 48/00 (2006.01)

A61P 21/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.01.2015 PCT/EP2015/051081**

87 Fecha y número de publicación internacional: **30.07.2015 WO15110449**

96 Fecha de presentación y número de la solicitud europea: **21.01.2015 E 15703234 (3)**

97 Fecha y número de publicación de la concesión europea: **16.12.2020 EP 3097197**

54 Título: **Elementos reguladores de ácido nucleico específicos de músculo y métodos y uso de los mismos**

30 Prioridad:

21.01.2014 EP 14151960

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
07.10.2021

73 Titular/es:

VRIJE UNIVERSITEIT BRUSSEL (33.3%)
Pleinlaan 2
1050 Brussel, BE;
VIB VZW (33.3%) y
UNIVERSITEIT GENT (33.3%)

72 Inventor/es:

VANDENDRIESSCHE, THIERRY;
CHUAH, MARINEE y
DE BLESER, PIETER

74 Agente/Representante:

ARIAS SANZ, Juan

ES 2 862 526 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Elementos reguladores de ácido nucleico específicos de músculo y métodos y uso de los mismos

5 **Campo**

La presente invención se refiere a elementos reguladores de ácido nucleico que son capaces de potenciar la expresión específica de músculo de genes, a métodos que emplean estos elementos reguladores y al uso de los mismos. La invención abarca además casetes de expresión, vectores y composiciones farmacéuticas que comprenden estos elementos reguladores. La presente invención es particularmente útil para aplicaciones que usan terapia génica, más particularmente, terapia génica dirigida al músculo y con fines de vacunación.

Antecedentes

15 El músculo es una diana atractiva para la terapia génica. La inserción de genes al músculo puede usarse para aumentar la expresión de proteínas estructurales musculares, tales como distrofina y sarcoglicanos, por ejemplo, para tratar la distrofia muscular. Además, el músculo puede usarse como plataforma terapéutica para expresar proteínas de la ruta reguladora/secretora no musculares para la diabetes, la aterosclerosis, la hemofilia, el cáncer, etc.

20 Los esfuerzos para insertar transgenes al músculo se han centrado en vectores derivados de adenovirus, retrovirus, lentivirus y virus adenoasociados (AAV), y plásmidos. Los vectores adenovirales tienen una capacidad de clonación relativamente grande que puede producirse a títulos altos y presentan una transducción relativamente eficaz del músculo. Desafortunadamente, estos vectores pueden provocar una respuesta inmunitaria celular robusta contra proteínas virales y algunas transgénicas. Además, pueden evocar una rápida activación del sistema inmunitario innato que puede contribuir a una respuesta inmunitaria inflamatoria limitante de la dosis y potencialmente peligrosa. Los vectores adenovirales no se integran en el genoma huésped, de modo que su capacidad para persistir durante largos períodos de tiempo no está clara. Los vectores retrovirales y lentivirales se integran de manera estable en el genoma celular diana, que permite potencialmente una transferencia génica persistente. Mientras que los vectores lentivirales pueden transducir tanto células que se dividen como que no se dividen, los vectores retrovirales convencionales derivados del virus de la leucemia murina de Moloney (MoMLV) solo pueden transducir células que se dividen. En consecuencia, pueden usarse vectores lentivirales para transducir células del músculo esquelético que no se dividen, mientras que éstas son refractarias a la transducción por inyección directa con vectores retrovirales. No obstante, incluso la transducción lentiviral del músculo esquelético no es muy eficiente. El ADN de plásmido desnudo muestra una notable capacidad para transferir genes al músculo. Los plásmidos muestran una inmunogenicidad y toxicidad mínimas, y tienen una capacidad de clonación extremadamente grande. La principal desventaja de los plásmidos es su eficiencia de transfección relativamente escasa bajo protocolos de inserción típicos. La retención de plásmidos es otra consideración importante.

40 El vector viral adenoasociado (AAV) es, con mucho, el vehículo de inserción génica más prometedor para la terapia génica dirigida al músculo. El tropismo natural del AAV a las células musculares, su expresión transgénica persistente a largo plazo, sus múltiples serotipos, así como su mínima respuesta inmunitaria han hecho que los vectores de AAV sean muy adecuados para la terapia génica dirigida al músculo. El vector de AAV puede insertarse tanto en el músculo esquelético como en el músculo cardíaco por medio de administraciones locales, regionales y sistémicas.

45 Sin embargo, sigue habiendo preocupaciones con respecto a la eficacia y seguridad de algunos enfoques de inserción génica. Los principales factores limitantes son: niveles de expresión transgénica transitorios y/o insuficientes y expresión inapropiada del transgén en tipos de células no deseadas. En particular, se ha demostrado que la expresión transgénica involuntaria en células presentadoras de antígeno (APC) aumenta el riesgo de respuestas inmunitarias adversas contra las células modificadas con genes y/o el producto transgénico terapéutico que, en consecuencia, reduce la expresión génica a largo plazo.

50 Los métodos convencionales de diseño de vectores se basaron en enfoques de prueba y error irregulares mediante los cuales se combinaron potenciadores transcripcionales con promotores para aumentar los niveles de expresión. Aunque esto a veces podía ser eficaz, a menudo daba como resultado combinaciones no productivas que daban como resultado niveles de expresión o bien modestos o bien no aumentados del gen de interés y/o pérdida de especificidad tisular. Además, estos enfoques convencionales no tuvieron en cuenta la importancia de incluir motivos reguladores conservados evolutivamente en los módulos de expresión, lo cual es particularmente relevante para la traducción clínica.

60 La solicitud de patente china CN103421786 da a conocer un plásmido de expresión que comprende la secuencia promotora del gen Mylpf de ratón de 1498 pb para dirigir la expresión de GFP. El plásmido se usó para transfectar células musculares de ratón C2D12. Sin embargo, el documento CN103421786 no se refiere a elementos reguladores pequeños, por ejemplo, de 1000 nucleótidos como máximo para potenciar la expresión génica específica de músculo. Katwal *et al.* (2013 Gene Therapy 20:930 -938) da a conocer el uso de un vector de AAV-9 que comprende el promotor de CK6, que incluye elementos reguladores de ese promotor, para restringir la expresión transgénica *in vivo* al músculo esquelético.

Los números de registro de Genbank AQ503558 y FI096286 dan a conocer secuencias de nucleótidos que parecen contener regiones idénticas a ciertas regiones de la secuencia promotora del gen *MyIpf* humano, pero no se ha atribuido ninguna función a dichas secuencias.

La solicitud de patente estadounidense US 2002/0198371 se refiere a la identificación y el mapeo de un número muy grande de polimorfismos de un solo nucleótido en todo el genoma humano.

Un enfoque computacional que depende de una matriz de diferencias de distancia modificada (DDM) (estrategia de escalado multidimensional (MDS)) (De Bleser *et al.* 2007. Genoma Biol 8, R83) ha demostrado ser útil para la identificación *in silico* de agrupaciones de motivos de sitio de unión a factores de transcripción (TFBS) conservados evolutivamente asociados con una expresión robusta específica de tejido en hígado (documento WO 2009/130208) y corazón (documento WO2011/051450).

Sigue existiendo la necesidad en la técnica de una inserción génica segura y eficaz al músculo.

Sumario

Los presentes inventores se han basado en una estrategia de DDM-MDS modificada (De Bleser *et al.*, 2007) combinada con una estrategia de cribado potenciada para identificar motivos de sitio de unión a factores de transcripción (TFBS) conservados evolutivamente asociados con genes específicos de músculo altamente expresados definidos en el presente documento como elementos reguladores de ácido nucleico. Tal como se muestra en la sección experimental, los inventores pudieron identificar elementos reguladores de ácido nucleico que potenciaban específicamente la expresión génica tanto en el músculo cardíaco como en el esquelético y elementos reguladores de ácido nucleico específicos de músculo esquelético. Estos elementos reguladores se validaron posteriormente *in vivo*, produciendo una expresión génica eficaz y específica de tejido. Por tanto, este enfoque permite el uso de dosis de vector más bajas y, por tanto, más seguras, al tiempo que maximiza la eficacia terapéutica.

La invención es tal como se define en las reivindicaciones. La divulgación proporciona los siguientes aspectos:

Aspecto 1: un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo, preferiblemente para potenciar la expresión génica específica de músculo esquelético, que comprende la secuencia de SEQ ID NO:10 y que tiene una longitud máxima de 550 nucleótidos, o que consiste en un fragmento funcional de dicha secuencia que tiene una longitud máxima de 300 nucleótidos y que comprende la secuencia de SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40 o SEQ ID NO:41.

Aspecto 2: el elemento regulador de ácido nucleico según el aspecto 1, en el que dicho fragmento funcional consiste en la secuencia de SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40 o SEQ ID NO:41.

Aspecto 3: uso *in vitro* o *ex vivo* del elemento regulador de ácido nucleico según el aspecto 1 o 2 en un casete de expresión de ácido nucleico, o un vector, más particularmente para potenciar la expresión específica de músculo de dicho casete o vector de expresión de ácido nucleico.

Aspecto 4: un elemento regulador de ácido nucleico según el aspecto 1 o 2 para su uso en medicina.

Aspecto 5: un casete de expresión de ácido nucleico que comprende al menos uno, tal como uno, dos, tres, cuatro, cinco o más, elemento regulador de ácido nucleico, operativamente unido a un promotor y un transgén, en el que dicho elemento regulador de ácido nucleico consiste en el elemento regulador de ácido nucleico tal como se define en la reivindicación 1 o 2.

Aspecto 6: el casete de expresión de ácido nucleico según el aspecto 5, en el que el promotor es un promotor específico de músculo, preferiblemente el promotor del gen de desmina (DES) o el promotor de SPc5-12.

Aspecto 7: el casete de expresión de ácido nucleico según uno cualquiera de los aspectos 5 o 6, en el que el transgén codifica una proteína terapéutica, una proteína inmunogénica o una proteína estructural, preferiblemente distrofina o sarcoglicano.

Aspecto 8: el casete de expresión de ácido nucleico según uno cualquiera de los aspectos 5 a 7, que comprende además un intrón, preferiblemente el intrón del Virus Diminuto del Ratón (MVM).

Aspecto 9: el casete de expresión de ácido nucleico según cualquiera de los aspectos 5 a 8, que comprende además una señal de poliadenilación, preferiblemente la señal de poliadenilación del virus de simios 40 (SV40) o el sitio de poliA sintética (SPA) tal como se describe en Levitt *et al.* (1989, Genes Dev 3:1019-1025) (SEQ ID NO:46).

Aspecto 10: un vector que comprende al menos un elemento regulador de ácido nucleico, o el casete de expresión de ácido nucleico según uno cualquiera de los aspectos 5 a 9, en el que dicho elemento regulador de ácido nucleico

consiste en el elemento regulador de ácido nucleico tal como se define en la reivindicación 1 o 2.

Aspecto 11: el vector según el aspecto 10, que es un vector viral, preferiblemente un vector viral adenoasociado (AAV), más preferiblemente un vector de AAV9.

Aspecto 12: el vector según el aspecto 10, que es un vector no viral, preferiblemente un plásmido, un minicírculo o un vector basado en transposón, tal como un vector basado en PiggyBac o un vector basado en Sleeping Beauty.

Aspecto 13: una composición farmacéutica que comprende el casete de expresión de ácido nucleico según uno cualquiera de los aspectos 5 a 9, o el vector según uno cualquiera de los aspectos 10 a 12, y un portador farmacéuticamente aceptable.

Aspecto 14: el elemento regulador de ácido nucleico según el aspecto 1 o 2, el casete de expresión de ácido nucleico según uno cualquiera de los aspectos 5 a 9, el vector según uno cualquiera de los aspectos 10 a 12, o la composición farmacéutica según el aspecto 13 para su uso en terapia génica, preferiblemente terapia génica dirigida al músculo.

Aspecto 15: el elemento regulador de ácido nucleico, el casete de expresión de ácido nucleico, el vector o la composición farmacéutica para su uso según el aspecto 14, en el que la terapia génica es para la distrofia muscular (por ejemplo, distrofia muscular de Duchenne (DMD)/distrofia muscular de Becker (DMO)), distrofia miotónica, enfermedad neuromuscular, enfermedades de neuronas motoras (MND), tales como por ejemplo enfermedad de Charcot-Marie-Tooth (CMT), atrofia muscular espinal (SMA), esclerosis lateral amiotrófica (ELA), distrofia muscular de Emery-Dreifuss, distrofia muscular facioescapulohumeral (FSHD), distrofias musculares congénitas, miopatías congénitas, distrofia muscular de la cintura y las extremidades, miopatías metabólicas, enfermedades inflamatorias musculares, miastenia, miopatías mitocondriales, anomalías de canales iónicos, enfermedades de la envoltura nuclear, cardiomiopatías, hipertrofia cardíaca, insuficiencia cardíaca, miopatías distales, enfermedades cardiovasculares, hemofilia, que incluye hemofilia A y B, y diabetes.

Aspecto 16: el elemento regulador de ácido nucleico según el aspecto 1 o 2, el casete de expresión de ácido nucleico según uno cualquiera de los aspectos 5 a 9, el vector según uno cualquiera de los aspectos 10 a 12, o la composición farmacéutica según el aspecto 13 para su uso como vacuna, preferiblemente una vacuna profiláctica, o para su uso en terapia de vacunación, preferiblemente vacunación profiláctica.

Aspecto 17: Un método *in vitro* o *ex vivo* para expresar un producto transgénico en células musculares, preferiblemente células de músculo cardíaco y/o células de músculo esquelético, que comprende:

- introducir el casete de expresión de ácido nucleico según uno cualquiera de los aspectos 5 a 9, o el vector según uno cualquiera de los aspectos 10 a 12 en las células musculares;

- expresar el producto transgénico en las células musculares.

Breve descripción de los dibujos

Figura 1: Diagrama de flujo de la identificación y validación de elementos reguladores de ácido nucleico. Se usó un enfoque computacional para identificar los elementos reguladores de ácido nucleico que implican las siguientes etapas: (1) identificación de genes específicos de tejido altamente expresados, por ejemplo, basándose en análisis estadístico de datos de expresión de micromatriz de tejidos humanos normales; (2) extracción de las secuencias de promotores correspondientes de bases de datos disponibles para el público; (3) identificación de los módulos reguladores y los sitios de unión a factores de transcripción (TFBS) que contienen; (4) a continuación, se buscó en el contexto genómico de los genes altamente expresados para agrupaciones conservadas evolutivamente de TFBS (es decir, elementos reguladores de ácido nucleico). Los elementos reguladores de ácido nucleico identificados se diseñaron *de novo* y se validaron *in vivo* sometiendo a prueba si la inclusión en un constructo aumenta la expresión de un gen indicador.

Figura 2: Secuencias de nucleótidos de elementos reguladores específicos de músculo cardíaco y esquelético CSk-SH1 (A, SEQ ID NO: 1), CSk-SH2 (B, SEQ ID NO: 2), CSk-SH3 (C, SEQ ID NO: 3), CSk-SH4 (D, SEQ ID NO: 4), CSk-SH5 (E, SEQ ID NO: 5) y CSk-SH6 (F, SEQ ID NO: 6).

La figura 3 muestra una representación esquemática de los vectores de AAV9sc-CSk-SH/Sk-SH-Des-Luc2 dados a conocer en el presente documento. El casete de expresión se empaquetó en un virus adenoasociado autocomplementario (sc), serotipo 9 (AAV9). El promotor de desmina (Des) específico de músculo cardíaco y esquelético regula la transcripción del transgén de luciferasa (Luc2). Los elementos reguladores de ácido nucleico específicos de músculo cardíaco y esquelético (CSk-SH) o específicos de músculo (Sk-SH) se clonaron en el sentido de 5' del promotor de Des. El casete de expresión comprende además el intrón del Virus Diminuto del Ratón (MVM) y una señal de poliadenilación (pA) del virus de simios 40 (SV40). El casete de expresión está flanqueado por repeticiones terminales invertidas (ITR) del virus adenoasociado, serotipo 2 (AAV2).

Figura 4: Diferencia en la expresión de luciferasa **(A)** y el nivel de ARNm de Luc **(B)** en el tejido cardíaco y muscular de ratones en los que se inyectó, por vía intravenosa, de izquierda a derecha, el vector de AAV9sc-CSk-SH1-Des-Luc2 (CSk-SH1), AAV9sc-CSk-SH2-Des-Luc2 (CSk-SH2), AAV9sc-CSk-SH3-Des-Luc2 (CSk-SH3), AAV9sc-CSk-SH4-Des-Luc2 (CSk-SH4), AAV9sc-CSk-SH5-Des-Luc2 (CSk-SH5) o AAV9sc-CSk-SH6-Des-Luc2 (CSk-SH6) en comparación con ratones en los que se inyectó el vector de control AAV9sc-Des-Luc2 sin elemento regulador de ácido nucleico (control, sin CSk-SH). La expresión de luciferasa se midió como flujo total, expresado en fotones por segundo por centímetro cuadrado por estereorradián (fotones/s/cm²/sr), liberados por la actividad luciferasa en los tejidos seleccionados a las 9 semanas después de la inyección. Los resultados se presentaron como media $\pm$ error estándar de la media, *p<0,05; **p<0,001. Los niveles de ARNm de Luc se midieron mediante un método de RT-PCR cuantitativa (qRT-PCR) al final del experimento a partir de ARN total extraído de biopsias de los tejidos indicados. Los resultados se presentaron como media $\pm$ error estándar de la media, *p<0,05; **p<0,001.

Figura 5: **(A)** Expresión de luciferasa en tejidos seleccionados de ratones en los que se inyectó, por vía intravenosa, de izquierda a derecha, el vector de AAV9sc-Des-Luc2 (control, sin CSk-SH), AAV9sc-CSk-SH1-Des-Luc2 (CSk-SH1), AAV9sc-CSk-SH2-Des-Luc2 (CSk-SH2), AAV9sc-CSk-SH3-Des-Luc2 (CSk-SH3), AAV9sc-CSk-SH4-Des-Luc2 (CSk-SH4), AAV9sc-CSk-SH5-Des-Luc2 (CSk-SH5) o AAV9sc-CSk-SH6-Des-Luc2 (CSk-SH6). La expresión de luciferasa se midió como flujo total, expresado en fotones por segundo por centímetro cuadrado por estereorradián (fotones/s/cm²/sr), liberados por la actividad luciferasa en el tejido seleccionado a las 9 semanas después de la inyección. Los resultados se presentaron como media $\pm$ error estándar de la media. **(B)** Nivel de ARNm de Luc en tejidos seleccionados de ratones en los que se inyectó, por vía intravenosa, de izquierda a derecha, el vector de AAV9sc-CSk-SH1-Des-Luc2 (CSk-SH1), AAV9sc-CSk-SH2-Des-Luc2 (CSk-SH2), AAV9sc-CSk-SH3-Des-Luc2 (CSk-SH3), AAV9sc-CSk-SH4-Des-Luc2 (CSk-SH4), AAV9sc-CSk-SH5-Des-Luc2 (CSk-SH5) o AAV9sc-CSk-SH6-Des-Luc2 (CSk-SH6) en comparación con ratones en los que se inyectó el vector de control AAV9sc-Des-Luc2 (control, sin CSk-SH). Los niveles de ARNm de Luc se midieron mediante un método de RT-PCR cuantitativa (qRT-PCR) al final del experimento a partir de ARN total extraído de biopsias de los tejidos indicados. Los resultados se presentaron como media $\pm$ error estándar de la media.

Figura 6: Eficiencia de transducción en diferentes órganos de ratones en los que se inyectó AAV9sc-CSk-SH1-Des-Luc2 (CSk-SH1) **(A)** o AAV9sc-CSk-SH5-Des-Luc2 (CSk-SH5) **(B)**. Se determinó el número de copias de AAV en relación con 100 ng de ADN genómico para ambos constructos a una dosis de 5x10⁹ vg/ratón.

Figura 7: Secuencias de nucleótidos de los elementos reguladores específicos de músculo Sk-SH1 (A, SEQ ID NO: 7), Sk-SH2 (B, SEQ ID NO: 8), Sk-SH3 (C, SEQ ID NO: 9), Sk-SH4 (D, SEQ ID NO: 10), Sk-SH5 (E, SEQ ID NO: 11), Sk-SH6 (F, SEQ ID NO: 12), Sk-SH7 (G, SEQ ID NO: 13).

La figura 8 muestra la diferencia en la expresión de luciferasa en tejido cardíaco y muscular de ratones en los que se inyectó, por vía intravenosa, de izquierda a derecha, el vector de AAV9sc-Sk-SH1-Des-Luc2 (Sk-SH1, n=4), AAV9sc-Sk-SH2-Des-Luc2 (Sk-SH2, n=4), AAV9sc-Sk-SH3-Des-Luc2 (Sk-SH3, n=4), AAV9sc-Sk-SH4-Des-Luc2 (Sk-SH4, n=4), AAV9sc-Sk-SH5-Des-Luc2 (Sk-SH5, n=1), AAV9sc-Sk-SH6-Des-Luc2 (Sk-SH5, n=3) o AAV9sc-Sk-SH7-Des-Luc2 (Sk-SH6, n=4) en comparación con ratones en los que se inyectó el vector de control AAV9sc-Des-Luc2 sin elemento regulador de ácido nucleico (control, sin Sk-SH, n=5). La expresión de luciferasa se midió como flujo total, expresado en fotones por segundo por centímetro cuadrado por estereorradián (fotones/s/cm²/sr), liberados por la actividad luciferasa en los tejidos seleccionados a las 7 semanas después de la inyección. Los resultados se presentaron como media $\pm$ error estándar de la media, *p<0,05; **p<0,001.

La figura 9 muestra la expresión de luciferasa en tejidos seleccionados de ratones en los que se inyectó por vía intravenosa, de izquierda a derecha, el vector de AAV9sc-Des-Luc2 (control, sin Sk-SH, n=5), AAV9sc-Sk-SH1-Des-Luc2 (Sk-SH1, n=4), AAV9sc-Sk-SH2-Des-Luc2 (Sk-SH2, n=4), AAV9sc-Sk-SH3-Des-Luc2 (Sk-SH3, n=4), AAV9sc-Sk-SH4-Des-Luc2 (Sk-SH4, n=4), AAV9sc-Sk-SH5-Des-Luc2 (Sk-SH5, n=1), AAV9sc-Sk-SH6-Des-Luc2 (Sk-SH5, n=3) o AAV9sc-Sk-SH7-Des-Luc2 (Sk-SH6, n=4). La expresión de luciferasa se midió como flujo total, expresado en fotones por segundo por centímetro cuadrado por estereorradián (fotones/s/cm²/sr), liberados por la actividad luciferasa en el tejido seleccionado a las 7 semanas después de la inyección. Los resultados se presentaron como media $\pm$ error estándar de la media, *p<0,05; **p<0,001.

Figura 10: **(A)** Diagrama esquemático del constructo de plásmido AAVsc-CMV-Luc2-SV40pA con indicación de dónde se clona el promotor de citomegalovirus (CMVp) en el sentido de 5' del gen de luciferasa de luciérnaga (Luc2). Las abreviaturas usadas son: ITR: repetición terminal invertida viral; SV40pA: sitio de poliadenilación del virus de simios. **(B)** Expresión de luciferasa en tejidos seleccionados de ratones en los que se inyectó, por vía intravenosa, de izquierda a derecha, el vector de AAV9sc-CMV-Luc2 (CMV, n=4), AAV9sc-Sk-SH4-Des-Luc2 (Sk-SH4, n=2), AAV9sc-CSk-SH1-Des-Luc2 (CSk-SH1, n=4) o AAV9sc-CSk-SH5-Des-Luc2 (CSk-SH5, n=4). La expresión de luciferasa se midió como flujo total, expresado en fotones por segundo por centímetro cuadrado por estereorradián (fotones/s/cm²/sr), liberados por la actividad luciferasa en el tejido seleccionado. Los resultados se presentaron como media $\pm$ error estándar de la media. **(C)** Diferencia en la expresión de luciferasa en tejido cardíaco y muscular de ratones en los que se inyectó, por vía intravenosa, de izquierda a derecha, el vector de AAV9sc-Sk-SH4-Des-Luc2 (Sk-SH4, n=2), AAV9sc-CSk-SH1-Des-Luc2 (CSk-SH1, n=4) o AAV9sc-CSk-SH5-Des-Luc2 (CSk-SH5, n=4) en

comparación con ratones en los que se inyectó el vector de AAV9sc-CMV-Luc2 (CMV, n=4). La expresión de luciferasa se midió como flujo total, expresado en fotones por segundo por centímetro cuadrado por estereorradián (fotones/s/cm²/sr), liberados por la actividad luciferasa en los tejidos seleccionados a las 7 semanas después de la inyección. Los resultados se presentaron como media $\pm$ error estándar de la media.

La figura 11 muestra fragmentos funcionales y los sitios de unión a factores de transcripción (TFBS) indicados de CSK-SH1 (**A**), CSK-SH-5 (**B**) y Sk-SH4 (**C**) mapeados en una representación esquemática de SEQ ID NO:1 (**A**), SEQ ID NO:5 (**B**) y SEQ ID NO:10 (**C**), respectivamente.

Figura 12: Diferencia en el nivel de ARNm de Luc en tejido cardíaco y muscular de ratones en los que se inyectó, por vía intravenosa, de izquierda a derecha, el vector de AAV9sc-Sk-SH1-Des-Luc2 (Sk-SH1), AAV9sc-Sk-SH2-Des-Luc2 (Sk-SH2), AAV9sc-Sk-SH3-Des-Luc2 (Sk-SH3), AAV9sc-Sk-SH4-Des-Luc2 (Sk-SH4), AAV9sc-Sk-SH6-Des-Luc2 (Sk-SH6) o AAV9sc-Sk-SH7-Des-Luc2 (Sk-SH7) en comparación con ratones en los que se inyectó el vector de control AAV9sc-Des-Luc2 sin elemento regulador de ácido nucleico (control, sin Sk-SH). Los niveles de ARNm de Luc se midieron mediante un método de RT-PCR cuantitativa (qRT-PCR) a partir de ARN total extraído de biopsias de los tejidos indicados. Los resultados se presentaron como media $\pm$ error estándar de la media, *p<0,05; **p<0,001.

Figura 13: Eficiencia de transducción en diferentes órganos de ratones en los que se inyectó el vector de AAV9sc-Sk-SH4-Des-MVM-Luc-pA (**A**) o el vector de AAV9sc-Sk-SH1-Des-MVM-Luc-pA (**B**). Se determinó el número de copias de AAV en relación con 100 ng de ADN genómico para ambos vectores (n=3).

La figura 14 muestra la expresión de luciferasa en tejido cardíaco y muscular de ratones en los que se inyectó, por vía intravenosa, de izquierda a derecha, el vector de AAV9sc-Des-Luc (control, sin Sk-SH), AAV9sc-Sk-SH4-Des-Luc (Sk-SH4), AAV9sc-Sk-SH4^a-Des-Luc (Sk-SH4^a), AAV9sc-Sk-SH4^b-Des-Luc (Sk-SH4^b), AAV9sc-Sk-SH4^c-Des-Luc (Sk-SH4^c), AAV9sc-Sk-SH4^d-Des-Luc (Sk-SH4^d) o AAV9sc-Sk-SH4^e-Des-Luc (Sk-SH4^e). La expresión de luciferasa se midió como flujo total, expresado en fotones por segundo por centímetro cuadrado por estereorradián (fotones/s/cm²/sr), liberados por la actividad luciferasa en el tejido seleccionado 5 semanas después de la inyección. Los resultados se presentaron como media $\pm$ error estándar de la media. Se indica la diferencia en veces en la expresión de luciferasa en ratones en los que se inyectó Sk-SH4 y Sk-SH4^b en comparación con ratones en los que se inyectó el vector de control AAV9sc-Des-Luc2 sin elemento regulador de ácido nucleico.

Figura 15: Ensayo de inmunoprecipitación de cromatina (CHIP) para tejido cardíaco (**A**, **C**) y muscular (**B**, **D**) de ratones en los que se inyectó AAV9sc-Sk-SH4-Des-Luc (**A**, **B**) o AAV9sc-CSK-SH5-Des-Luc (**C**, **D**) (5x10⁹ vg/ratón). Se usaron anticuerpos específicos para los factores de transcripción CEBP, SRF y MEF2. Se diseñaron cebadores de PCR para amplificar una región de Sk-SH4 que se une a CEBP y SRF, o una región de CSK-SH5 que se une a CEBP y MEF2, y como control negativo (-) se usó una región no transcrita en el cromosoma 17. Se determinaron los acontecimientos de unión para 10³ células para cada uno de los pares de cebadores correspondientes. Los resultados se presentan como media $\pm$ error estándar de la media. Se indica una diferencia significativa en comparación con el control negativo (prueba de la t, * P≤0,05).

La figura 16 muestra representaciones esquemáticas de constructos de plásmidos de AAV monocatenarios (*single-stranded*, ss) dados a conocer en el presente documento, que comprenden el transgén de microdistrofina1 (MD1) (**A**, AAVss-SkSH4-Des-MVM-MD1) o folistatina (FST) (**C**, AAVssSkSH4-Des-MVM-FST-2A-Luc2) regulado por el promotor de desmina operativamente unido al elemento regulador de ácido nucleico específico de músculo SkSH4 clonado en el sentido de 5' del promotor de desmina. El gen de folistatina se unió al gen indicador de Luc2 a través de un péptido 2A. Los casetes de expresión comprenden además el intrón del Virus Diminuto del Ratón (MVM) y el sitio de poliadenilación (pA) sintético Proudfoot de 49 pb. Los casetes de expresión están flanqueados por repeticiones terminales invertidas (ITR). (**B**) Secuencia de nucleótidos del constructo de plásmido AAVss-SkSH4-Des-MVM-MD1 (SEQ ID NO:44). (**D**) Secuencia de nucleótidos del constructo de plásmido AAVss-SkSH4-Des-MVM-FST-2A-Luc2 (SEQ ID NO:45).

Figura 17: Prueba de cinta de correr para ratones MDX-SCID en los que se inyectó, de izquierda a derecha, PBS (control), AAVss-Sk-SH4-Des-MVMFST-2A-Luc, AAVss-Sk-SH4-Des-MVM-MD1, o la combinación de AAVss-Sk-SH4-Des-MVM-FST-2A-Luc y AAVss-Sk-SH4-Des-MVM-MD1. Los resultados se expresan como la distancia calculada cubierta por cada grupo de ratones cuando corren en una máquina de cinta de correr.

Figura 18: (**A**) Tinción con hematoxilina/eosina de tejidos del músculo gastrocnemio de ratones MDX/SCID en los que se inyectó PBS (**a**, control), AAVss-Sk-SH4-Des-MVM-FST-2A-Luc (**b**), AAVss-Sk-SH4-Des-MVM-MD1 (**c**) o la combinación de AAVss-Sk-SH4-Des-MVM-MD1 y AAVss-Sk-SH4-Des-MVM-FST-2A-Luc (**d**). (**B**) Cuantificación de células nucleadas centralmente en ratones silvestres C57Bl/6, ratones MDX/SCID no tratados (control) y ratones MDX/SCID en los que se inyectó AAVss-Sk-SH4-Des-MVM-MD1 (MD1), AAVss-Sk-SH4-Des-MVM-FST-2A-Luc (FST) o ambos vectores terapéuticos (FST+MD1). El análisis estadístico se realizó sobre miofibras transectadas transversalmente teñidas con H&E del músculo gastrocnemio incrustado en parafina. *** ≤0,0001 **≤0,001 *≤0,05

Figura 19: Niveles de ARNm de microdistrofina1 (MD1) (**A**, **B**) y folistatina (FST) (**C**) en tejidos cardíacos y musculares (gastrocnemio y cuádriceps) de ratones en los que se inyectó, por vía intravenosa, el vector de AAVss-Sk-SH4-Des-

MVM-MD1 (A), la combinación de AAVss-Sk-SH4-Des-MVM-MD1 y AAVss-Sk-SH4-Des-MVM-FST-2A-Luc (B), o AAVss-Sk-SH4-Des-MVM-FST (C) en relación con la expresión del gen de mantenimiento endógeno (GAPDH: gliceraldehído-3-fosfato deshidrogenasa). Los resultados se presentan como expresión relativa de MD1 o FST ($\Delta\Delta CT$).

La figura 20 muestra representaciones esquemáticas de constructos de AAV autocomplementarios (*self-complementary*, sc) dados a conocer en el presente documento, que comprenden el transgén de luciferasa (Luc) regulado por (a) el promotor de desmina (desmina) específico de músculo cardíaco y esquelético, (b) el promotor de SPc5-12, (c) el promotor de citomegalovirus (CMV), (d) el promotor de desmina operativamente unido a un elemento regulador de ácido nucleico específico de músculo (Sk-SH) clonado en el sentido de 5' del promotor de desmina, o (e) el promotor de SPc5-12 operativamente unido a un elemento regulador de ácido nucleico específico de músculo (Sk-SH) clonado en el sentido de 5' del promotor de SPc5-12. Los casetes de expresión (a), (b), (d) y (e) comprenden además el intrón del Virus Diminuto del Ratón (MVM) y una señal de poliadenilación (pA). Los casetes de expresión están flanqueados por repeticiones terminales invertidas (ITR).

Figura 21: Diferencia en los niveles de ARNm de Luc en tejidos seleccionados de ratones CB17/lcrTac/Prkdcscid en los que se inyectó, por vía intravenosa, de izquierda a derecha, el vector de AAVsc-CMV-Luc, AAVsc-Des-MVM-Luc, AAVsc-Sk-SH4-Des-MVM-Luc, AAVsc-SPc5-12-MVM-Luc o AAVsc-Sk-SH4-SPc5-12-MVM-Luc tal como se muestra de manera esquemática en la figura 20, en comparación con ratones en los que se inyectó el vector de AAVsc-CMV-Luc. Los niveles de ARNm de Luc se midieron mediante qRT-PCR a partir de ARN total extraído de biopsias de los tejidos indicados. Los resultados se presentaron como media $\pm$ error estándar de la media, * $p < 0,05$; ** $p < 0,001$.

Descripción

Tal como se usan en el presente documento, las formas singulares “uno”, “una” y “el/la” incluyen referencias tanto singulares como plurales, a menos que el contexto indique claramente lo contrario.

Los términos “que comprende”, “comprende” y “compuesto por” tal como se usan en el presente documento son sinónimos de “que incluye”, “incluye” o “que contiene”, “contiene”, y son inclusivos o de extremos abiertos y no excluyen miembros, elementos o etapas de método adicionales no mencionados. Los términos también abarcan “que consiste en” y “que consiste esencialmente en”, que gozan de significados bien establecidos en la terminología de patentes.

La mención de intervalos numéricos por puntos extremos incluye todos los números y fracciones incluidos dentro de los respectivos intervalos, así como los puntos extremos mencionados.

Los términos “alrededor” o “aproximadamente” tal como se usan en el presente documento cuando se refieren a un valor medible, tal como un parámetro, una cantidad, una duración temporal, etc., pretenden abarcar variaciones de y desde el valor especificado, tales como variaciones de $\pm 10\%$ o menos, preferiblemente $\pm 5\%$ o menos, más preferiblemente $\pm 1\%$ o menos, y aún más preferiblemente $\pm 0,1\%$ o menos del valor especificado, en la medida en que tales variaciones sean apropiadas de realizar en la invención dada a conocer. Debe entenderse que el valor al que se refiere el modificador “aproximadamente” en sí mismo también se da a conocer específica y preferiblemente.

Mientras que los términos “uno o más” o “al menos uno”, tal como uno o más miembros o al menos un miembro de un grupo de miembros, están claros per se, por medio de ejemplificación adicional, el término abarca, entre otras cosas, una referencia a uno cualquiera de dichos miembros, o a dos o más de dichos miembros, tales como, por ejemplo, cualquier ≥ 3 , ≥ 4 , ≥ 5 , ≥ 6 o ≥ 7 , etc. de dichos miembros, y hasta todos de dichos miembros. En otro ejemplo, “uno o más” o “al menos uno” puede referirse a 1, 2, 3, 4, 5, 6, 7 o más.

Se incluye la descripción de los antecedentes de la invención en el presente documento para explicar el contexto de la invención. Esto no debe tomarse como una admisión de que alguno de los materiales a los que se hace referencia se publicó, se conoció o formaba parte del conocimiento general común en algún país a partir de la fecha de prioridad de cualquiera de las reivindicaciones.

Salvo que se defina otra cosa, todos los términos usados en la divulgación de la invención, incluyendo los términos técnicos y científicos, tienen el significado comúnmente entendido por un experto habitual en la técnica a la que pertenece esta invención. Por medio de orientación adicional, se incluyen definiciones de términos para apreciar mejor la enseñanza de la invención. Cuando se definen términos específicos en conexión con un aspecto particular de la invención o una realización particular de la invención, tal connotación se aplica a lo largo de esta memoria descriptiva, es decir, también en el contexto de otros aspectos o realizaciones de la invención, a menos que se defina otra cosa.

En los siguientes pasajes se definen con más detalle diferentes aspectos o realizaciones de la invención. Cada aspecto o realización así definido puede combinarse con otro(s) aspecto(s) o realización/realizaciones cualquiera/cualesquiera a menos que se indique claramente lo contrario. En particular, cualquier rasgo indicado como preferido o ventajoso puede combinarse con cualquier otra característica o rasgo indicado como preferido o ventajoso.

La referencia a lo largo de esta memoria descriptiva a “una realización” significa que un rasgo, estructura o característica particular descrita en conexión con la realización se incluye en al menos una realización de la presente

invención. Por tanto, las apariciones de la frase “en una realización” en diversos lugares a lo largo de esta memoria descriptiva no necesariamente todas hacen referencia a la misma realización, pero puede ocurrir. Además, los rasgos, estructuras o características particulares pueden combinarse de cualquier manera adecuada, tal como resultaría evidente para un experto en la técnica de esta divulgación, en una o más realizaciones. Además, mientras que algunas realizaciones descritas en el presente documento incluyen algunos pero no otros rasgos incluidos en otras realizaciones, las combinaciones de rasgos de diferentes realizaciones están destinadas a estar dentro del alcance de la invención, y formar diferentes realizaciones, tal como entenderían los expertos en la técnica. Por ejemplo, en las reivindicaciones adjuntas, cualquiera de las realizaciones reivindicadas puede usarse en cualquier combinación.

Para métodos generales relacionados con la invención, se hace referencia, entre otras cosas, a libros de texto muy conocidos, tales como, por ejemplo, “Molecular Cloning: A Laboratory Manual, 2ª ed.” (Sambrook *et al.*, 1989), “Current Protocols in Molecular Biology” (Ausubel *et al.*, 1987).

En el presente documento se da a conocer un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo que comprende, que consiste esencialmente en (es decir, el elemento regulador puede comprender por ejemplo adicionalmente secuencias usadas para fines de clonación, pero las secuencias indicadas constituyen la parte esencial del elemento regulador, por ejemplo, no forman parte de una región reguladora más grande tal como un promotor), o que consiste en una secuencia seleccionada del grupo que consiste en: SEQ ID NO:1, SEQ ID NO: 2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, o un fragmento funcional de las mismas (es decir, un fragmento funcional de una secuencia seleccionada del grupo que consiste en: SEQ ID NO:1, SEQ ID NO: 2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, o de una secuencia que tiene un alto porcentaje de identidad de secuencia con cualquiera de dichas secuencias).

Un ‘elemento regulador’ tal como se usa en el presente documento se refiere a un elemento de control transcripcional, en particular un elemento de control transcripcional de actuación en cis no codificante, capaz de regular y/o controlar la transcripción de un gen, en particular la transcripción específica de tejido de un gen. Los elementos reglamentarios comprenden al menos un sitio de unión a factores de transcripción (TFBS), más en particular al menos un sitio de unión para un factor de transcripción específico de tejido, lo más particularmente al menos un sitio de unión para un factor de transcripción específico de músculo. Normalmente, los elementos reguladores usados en el presente documento aumentan o potencian la expresión dirigida por promotores en comparación con la transcripción del gen a partir del promotor solo, sin los elementos reguladores. Por tanto, los elementos reguladores comprenden particularmente secuencias potenciadoras, aunque debe entenderse que los elementos reguladores que potencian la transcripción no se limitan a secuencias potenciadoras en el sentido de 5' lejanas típicas, sino que pueden aparecer a cualquier distancia del gen que regulan. De hecho, en la técnica se conoce que las secuencias que regulan la transcripción pueden estar situadas o bien en el sentido de 5' (por ejemplo, en la región promotora) o bien en el sentido de 3' (por ejemplo, en la UTR en 3) del gen que regulan *in vivo*, y pueden ubicarse en las inmediaciones del gen o más lejos. Cabe destacar que, aunque los elementos reguladores dados a conocer en el presente documento normalmente comprenden secuencias que se producen de manera natural, combinaciones de (partes de) tales elementos reguladores o varias copias de un elemento regulador, es decir, elementos reguladores que comprenden secuencias que no se producen de manera natural, se consideran también por sí mismos como elemento regulador. Los elementos reguladores tal como se usan en el presente documento pueden comprender parte de una secuencia más grande implicada en el control transcripcional, por ejemplo, parte de una secuencia promotora. Sin embargo, los elementos reguladores solos no son normalmente suficientes para iniciar la transcripción, sino que requieren un promotor para este fin. Los elementos reguladores dados a conocer en el presente documento se proporcionan como moléculas de ácido nucleico, es decir, ácidos nucleicos aislados o moléculas de ácido nucleico aisladas. Dicho elemento regulador de ácido nucleico tiene por tanto una secuencia que es solo una pequeña parte de la secuencia genómica que se produce de manera natural y por tanto no se produce de manera natural como tal, sino que está aislada de la misma.

El término “ácido nucleico” tal como se usa en el presente documento se refiere normalmente a un oligómero o polímero (preferiblemente un polímero lineal) de cualquier longitud compuesto esencialmente por nucleótidos. Una unidad de nucleótido incluye comúnmente una base heterocíclica, un grupo azúcar y al menos uno, por ejemplo, uno, dos o tres grupos fosfato, incluyendo grupos fosfato modificados o sustituidos. Las bases heterocíclicas pueden incluir, entre otras, bases de purina y pirimidina, tales como adenina (A), guanina (G), citosina (C), timina (T) y uracilo (U) que están muy extendidas en ácidos nucleicos que se producen de manera natural, otras bases que se producen de manera natural (por ejemplo, xantina, inosina, hipoxantina), así como bases química o bioquímicamente modificadas (por ejemplo, metiladas), no naturales o derivatizadas. Los grupos azúcar pueden incluir, entre otros, grupos pentosa (pentofurano), tales como preferiblemente ribosa y/o 2-desoxirribosa común en ácidos nucleicos que se producen de manera natural, o grupos azúcar de arabinosa, 2-desoxiarabinosa, treosa o hexosa, así como grupos azúcar modificados o sustituidos. Los ácidos nucleicos tal como se pretende en el presente documento pueden incluir nucleótidos naturales, nucleótidos modificados o mezclas de los mismos. Un nucleótido modificado puede incluir una base heterocíclica modificada, un resto azúcar modificado, un grupo fosfato modificado o una combinación de los

5 mismos. Pueden introducirse modificaciones de grupos fosfato o azúcares para mejorar la estabilidad, resistencia a la degradación enzimática o alguna otra propiedad útil. El término "ácido nucleico" abarca además preferiblemente moléculas de ADN, ARN e híbridas de ADN/ARN, incluyendo específicamente ARNhn, pre-ARNm, ARNm, ADNc, ADN genómico, productos de amplificación, oligonucleótidos e híbridos de ADN/ARN, ARN o ADN sintéticos (por ejemplo, sintetizados químicamente). Un ácido nucleico puede producirse de manera natural, por ejemplo, presente en o aislado de la naturaleza; o puede producirse de manera no natural, por ejemplo, recombinante, es decir, producirse por tecnología de ADN recombinante, y/o química o bioquímicamente, de manera parcial o completa. Un "ácido nucleico" puede ser bicatenario, parcialmente bicatenario o monocatenario. Cuando es monocatenario, el ácido nucleico puede ser la hebra sentido o la hebra antisentido. Además, el ácido nucleico puede ser circular o lineal.

10 Tal como se usa en el presente documento, "sitio de unión a factor de transcripción", "secuencia de unión a factor de transcripción" o "TFBS" se refiere a una secuencia de una región de ácido nucleico a la que se unen factores de transcripción. Los ejemplos no limitativos de TFBS incluyen sitios de unión para el factor de transcripción 3, también conocido como TCF3 o E2A; sitios de unión para el factor nuclear I, también conocido como NF1; sitios de unión para la proteína de unión a potenciador-CCAAT, también conocido como C/EBP; sitios de unión para diferenciación miogénica, también conocido como MyoD; sitios de unión para la proteína de unión a elementos reguladores de esteroides, también conocidos como SREBP; sitios de unión para el factor relacionado con leucemia/linfoma, también conocido como LRF; sitios de unión para la proteína 53, también conocida como p53; sitios de unión para el factor nuclear de hepatocitos 3-alfa, también conocido como HNF3a; sitios de unión para el factor nuclear de hepatocitos 3-beta, también conocido como HNF3b; sitios de unión para el factor nuclear de hepatocitos 4, también conocido como HNF4; sitios de unión para el factor potenciador específico de miocitos 2A, también conocido como MEF2A o RSRFC4; sitios de unión para el receptor activado por proliferador de peroxisoma, también conocido como PPAR; sitios de unión para el factor de respuesta sérica, también conocido como SRF; sitios de unión para la proteína 1b similar a activador de transcripción, también conocido como Tal1_b. Pueden encontrarse sitios de unión a factores de transcripción en bases de datos tales como Transfac®.

15 Las secuencias dadas a conocer en el presente documento pueden formar parte de secuencias de elementos reguladores capaces de controlar la transcripción de genes específicos de músculo *in vivo*, en particular controlar los siguientes genes: desmina también conocida como DES, CSM1 o CSM2; actinina, alfa 2 también conocida como ACTN2 o CMD1AA; filamina C (FLNC) también conocida como proteína similar a actina (ABLP), filamina-2 (FLN2), ABP-280, ABP280A, ABPA, ABPL, MFM5 o MPD4; ATPasa 1 de calcio del retículo sarcoplásmico/endoplasmático también conocido como ATP2A1, ATP2A o SERCA1; troponina I de tipo 1 (músculo esquelético lento) también conocida como TNNI1, SSTNI o TNN1; cadena ligera de miosina de músculo esquelético rápido fosforilable (MYLPF); miosina-1 también conocida como MYH1, MYHSA1, MYHa; MyHC-2X/D, o MyHC-2x; cadena alfa-3 de tropomiosina también conocida como TPM3, CFTD, NEM1, OK/SW-cl.5, TM-5, TM3, TM30, TM30nm, TM5, TPMsk3, TRK, hTM5 o hscp30; y proteína 2 que contiene dominios de repetición de anquirina también conocida como ANKRD2, o ARPP. En consecuencia, los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden comprender una secuencia de elementos reguladores de Des, es decir, elementos reguladores que controlan la expresión del gen Des *in vivo*, por ejemplo, elementos reguladores que comprenden SEQ ID NO:1, SEQ ID NO:17, SEQ ID NO:2, SEQ ID NO:18, o fragmentos funcionales de los mismos. Los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden comprender una secuencia de elementos reguladores de ACTN2, es decir, elementos reguladores que controlan la expresión *in vivo* del gen ACTN2, por ejemplo elementos reguladores que comprenden SEQ ID NO:3, SEQ ID NO:19, SEQ ID NO:4, SEQ ID NO:20, o fragmentos funcionales de los mismos. Los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden comprender una secuencia de elementos reguladores de FLNC, es decir, elementos reguladores que controlan la expresión *in vivo* del gen FLNC, por ejemplo elementos reguladores que comprenden SEQ ID NO:5, SEQ ID NO:21, SEQ ID NO:6, SEQ ID NO:22, o fragmentos funcionales de los mismos. Los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden comprender una secuencia de elementos reguladores de ATP2A1, es decir, elementos reguladores que controlan la expresión *in vivo* del gen ATP2A1, por ejemplo elementos reguladores que comprenden SEQ ID NO:7, SEQ ID NO:23, o fragmentos funcionales de los mismos. Los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden comprender una secuencia a partir de elementos reguladores TNNI1, es decir, elementos reguladores que controlan la expresión *in vivo* del gen TNNI1, por ejemplo elementos reguladores que comprenden SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:9, SEQ ID NO:25, o fragmentos funcionales de los mismos. Los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden comprender una secuencia de elementos reguladores de MYLPF, es decir, elementos reguladores que controlan la expresión del gen MYLPF *in vivo*, por ejemplo, elementos reguladores que comprenden SEQ ID NO:10, SEQ ID NO:26, o fragmentos funcionales de los mismos, tales como elementos reguladores que comprenden o que consisten en SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40 o SEQ ID NO:41. Los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden comprender una secuencia de elementos reguladores de MYH1, es decir, elementos reguladores que controlan la expresión del gen MYH1 *in vivo*, por ejemplo elementos reguladores que comprenden SEQ ID NO:11, SEQ ID NO:27, o fragmentos funcionales de los mismos. Los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden comprender una secuencia de elementos reguladores de TPM3, es decir, elementos reguladores que controlan la expresión *in vivo* del gen TPM3, por ejemplo elementos reguladores que comprenden SEQ ID NO:12, SEQ ID NO:28, o fragmentos funcionales de los mismos. Los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden comprender una secuencia de elementos reguladores de ANKRD2, es decir, elementos reguladores que controlan la expresión del

gen ANKRD2 *in vivo*, por ejemplo elementos reguladores que comprenden SEQ ID NO:13, SEQ ID NO:29, o fragmentos funcionales de los mismos.

Tal como se usa en el presente documento, los términos "identidad" e "idéntico" y similares se refieren a la similitud de secuencia entre dos moléculas poliméricas, por ejemplo, entre dos moléculas de ácido nucleico, por ejemplo, dos moléculas de ADN. Pueden hacerse alineaciones de secuencias y la determinación de la identidad de secuencia, por ejemplo, usando la herramienta de búsqueda de alineación local básica (BLAST) originalmente descrita por Altschul *et al.* 1990 (J Mol Biol 215: 403-10), tal como el algoritmo "Blast 2 sequences" descrito por Tatusova y Madden 1999 (FENS Microbiol Lett 174: 247-250). Normalmente, el porcentaje de identidad de secuencia se calcula a lo largo de toda la longitud de la secuencia. Tal como se usa en el presente documento, el término "sustancialmente idéntico" indica al menos el 90 %, preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 %, de identidad de secuencia.

El término 'fragmento funcional' tal como se usa en la solicitud se refiere a fragmentos de las secuencias de elementos reguladores dados a conocer en el presente documento que conservan la capacidad de regular la expresión específica de músculo, es decir, todavía pueden conferir especificidad de tejido y son capaces de regular expresión de un (trans)gen del mismo modo (aunque posiblemente no en la misma medida) que la secuencia de la que se derivan. Los fragmentos funcionales pueden comprender preferiblemente al menos 20, al menos 25, al menos 30, al menos 35, al menos 40, al menos 45, al menos 50, al menos 60, al menos 70, al menos 80, al menos 90, al menos 100, al menos 120, al menos 150, al menos 200, al menos 250, al menos 300, al menos 350, o al menos 400 nucleótidos contiguos de la secuencia de la que se derivan. También preferiblemente, los fragmentos funcionales pueden comprender al menos 1, más preferiblemente al menos 2, al menos 3, o al menos 4, incluso más preferiblemente al menos 5, al menos 10 o al menos 15 de los sitios de unión a factores de transcripción (TFBS) que están presentes en la secuencia de la que se derivan.

"Expresión específica de músculo" tal como se usa en la solicitud se refiere a la expresión preferente o predominante de un (trans)gen (tal como ARN y/o polipéptido) en músculos o tejidos musculares, en comparación con otros tejidos (es decir, no musculares). Al menos el 50 %, más particularmente al menos el 60 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 %, al menos el 98 %, al menos el 99 % o el 100 % de la expresión (trans)génica puede producirse dentro del músculo. La expresión específica de músculo puede suponer que no haya 'fuga' de producto genético expresado a otros órganos o tejidos distintos del músculo, tales como pulmón, hígado, cerebro, riñón y/o bazo.

Tal como se usa en el presente documento, "expresión específica del músculo cardíaco y esquelético" se refiere a la expresión preferente o predominante de un (trans)gen en el corazón, en particular el músculo cardíaco, y el músculo esquelético. Al menos el 50 %, más particularmente al menos el 60 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 %, al menos el 98 %, al menos el 99 % o el 100 % de la expresión (trans)génica puede producirse dentro del corazón y el músculo esquelético. Por tanto, menos del 10 %, menos del 5 %, menos del 2 % o incluso menos del 1 % de la expresión (trans)génica puede producirse en un órgano o tejido distinto del corazón y el músculo esquelético.

Tal como se usa en el presente documento, "expresión específica del músculo esquelético" se refiere a la expresión preferente o predominante de un (trans)gen en el músculo esquelético. Al menos el 50 %, más particularmente al menos el 60 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 %, al menos el 98 %, al menos el 99 % o el 100 % de la expresión (trans)génica puede producirse dentro del músculo esquelético. Por tanto, menos del 10 %, menos del 5 %, menos del 2 % o incluso menos del 1 % de la expresión (trans)génica puede producirse en un órgano o tejido distinto del músculo esquelético.

Lo mismo se aplica cambiando lo que se tenga que cambiar para la expresión específica de miocitos y específica de mioblastos, que puede considerarse como una forma particular de expresión específica de músculo. A lo largo de toda la solicitud, cuando se menciona específico de músculo en el contexto de la expresión, también se contemplan explícitamente expresión específica de miocitos y expresión específica de mioblastos. De manera similar, cuando se usa expresión específica de músculo cardíaco y esquelético en la solicitud, también se prevé explícitamente la expresión específica de miocitos esqueléticos y cardiomiocitos y la expresión específica de mioblastos esqueléticos y mioblastos cardíacos. De manera similar, cuando se usa expresión específica del músculo esquelético en la solicitud, también se prevé explícitamente la expresión específica de miocitos esqueléticos y la expresión específica de mioblastos esqueléticos.

Tal como se usan en el presente documento, los términos "músculo del corazón" o "músculo cardíaco" se refieren al tipo de músculo estriado regulado de forma autónoma que se encuentra en el corazón.

Tal como se usa en el presente documento, el término "músculo esquelético" se refiere al tipo de músculo estriado y controlado voluntariamente que está unido al esqueleto. Los ejemplos no limitativos de músculo esquelético incluyen el bíceps, el tríceps, el cuádriceps, el tibial interior y el músculo gastrocnemio.

El término "miocito", tal como se usa en el presente documento, se refiere a una célula que se ha diferenciado de un

mioblasto progenitor de manera que es capaz de expresar un fenotipo específico del músculo en condiciones apropiadas. Los miocitos diferenciados terminalmente se fusionan entre sí para formar miotubos, un componente importante de las fibras musculares. El término "miocito" también se refiere a miocitos que están desdiferenciados. El término incluye células *in vivo* y células cultivadas *ex vivo*, independientemente de si tales células son primarias o se han sometido a pases.

El término "mioblasto" usado en el presente documento se refiere a una célula embrionaria en el mesodermo que se diferencia para dar lugar a una célula muscular o miocito. El término incluye células *in vivo* y células cultivadas *ex vivo*, independientemente de si tales células son primarias o se han sometido a pases.

En el presente documento se da a conocer un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo que comprende, que consiste esencialmente en o que consiste en un fragmento funcional de una secuencia seleccionada del grupo que consiste en: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, en el que dicho fragmento funcional comprende al menos 20, preferiblemente al menos 25, más preferiblemente al menos 50, al menos 100, al menos 200 o al menos 250 nucleótidos contiguos de la secuencia de la que se deriva, y en el que dicho fragmento funcional comprende al menos 1, preferiblemente al menos 5, más preferiblemente al menos 10 o al menos 15 de los sitios de unión a factores de transcripción (TFBS) que están presentes en la secuencia de la que se deriva.

Un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo cardíaco y esquelético puede comprender un fragmento funcional de una secuencia seleccionada del grupo que consiste en: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, o una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, en el que dicho fragmento funcional comprende al menos 20, preferiblemente al menos 25, más preferiblemente al menos 50, al menos 100, al menos 200 o al menos 250 nucleótidos contiguos de la secuencia de la que se deriva, y en el que dicho fragmento funcional comprende al menos 1, preferiblemente al menos 5, más preferiblemente al menos 10 o al menos 15 de los sitios de unión a factores de transcripción (TFBS) que están presentes en la secuencia de la que se deriva.

Más particularmente, un elemento regulador de ácido nucleico para mejorar la expresión génica específica de músculo cardíaco y esquelético puede comprender un fragmento funcional de una secuencia seleccionada del grupo que consiste en: SEQ ID NO:1 y SEQ ID NO:5, o una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, en el que dicho fragmento funcional consiste en la secuencia de nucleótidos desde la posición 33 hasta 58 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 90 hasta 142 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 143 hasta 233 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 240 hasta 310 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 90 hasta 233 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 47 hasta 130 en SEQ ID NO:5; la secuencia de nucleótidos desde la posición 252 hasta 293 en SEQ ID NO:5; o la secuencia de nucleótidos desde la posición 330 hasta 450 en SEQ ID NO:5.

Un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo esquelético puede comprender un fragmento funcional de una secuencia seleccionada del grupo que consiste en: SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, en el que dicho fragmento funcional comprende al menos 20, preferiblemente al menos 25, más preferiblemente al menos 50, al menos 100, al menos 200 o al menos 250 nucleótidos contiguos de la secuencia de la que se deriva, y en el que dicho fragmento funcional comprende al menos 1, preferiblemente al menos 5, más preferiblemente al menos 10 o al menos 15 de los sitios de unión a factores de transcripción (TFBS) presentes en la secuencia de la que se deriva.

Más particularmente, un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo esquelético puede comprender un fragmento funcional de SEQ ID NO:10 o una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con SEQ ID NO:10, en el que dicho fragmento funcional consiste en la secuencia de nucleótidos desde la posición 10 hasta 180 en SEQ ID NO:10; la secuencia de nucleótidos desde la posición 190 hasta 240 en SEQ ID NO:10; la secuencia de nucleótidos desde la posición 241 hasta 300 en SEQ ID NO:10; la secuencia de nucleótidos desde la posición 241 hasta 360 en SEQ ID NO:10; o la secuencia de nucleótidos desde la posición 380 hasta 420 en SEQ ID NO:10.

Por ejemplo, los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden

comprender o consistir en una secuencia seleccionada del grupo que consiste en: la secuencia de nucleótidos desde la posición 33 hasta 58 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 90 hasta 142 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 143 hasta 233 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 240 hasta 310 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 90 hasta 233 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 47 hasta 130 en SEQ ID NO:5; la secuencia de nucleótidos desde la posición 252 hasta 293 en SEQ ID NO:5; la secuencia de nucleótidos desde la posición 330 hasta 450 en SEQ ID NO:5; la secuencia de nucleótidos desde la posición 10 hasta 180 en SEQ ID NO:10; la secuencia de nucleótidos desde la posición 190 hasta 240 en SEQ ID NO:10; la secuencia de nucleótidos desde la posición 241 hasta 300 en SEQ ID NO:10; la secuencia de nucleótidos desde la posición 241 hasta 360 en SEQ ID NO:10; la secuencia de nucleótidos desde la posición 380 hasta 420 en SEQ ID NO:10; o una secuencia que tiene al menos el 95 % de identidad con cualquiera de dichas secuencias.

Por ejemplo, los elementos reguladores de ácido nucleico dados a conocer en el presente documento comprenden o consisten en una secuencia seleccionada del grupo que consiste en: la secuencia de nucleótidos desde la posición 33 hasta 310 en SEQ ID NO:1; la secuencia de nucleótidos desde la posición 47 hasta 450 en SEQ ID NO:5; la secuencia de nucleótidos desde la posición 10 hasta 420 en SEQ ID NO:10, o una secuencia que tiene al menos el 95 % de identidad con cualquiera de dichas secuencias.

Además, se da a conocer en el presente documento un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo que comprende, que consiste esencialmente en o consiste en una secuencia seleccionada del grupo que consiste en: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, o una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias.

Un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo cardíaco y esquelético puede comprender una secuencia seleccionada del grupo que consiste en: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, o una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias.

Un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo esquelético puede comprender, consistir esencialmente en o consistir en una secuencia seleccionada del grupo que consiste en: SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, o una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias.

Se dan a conocer además en el presente documento elementos reguladores de ácido nucleico para potenciar la expresión génica específica de músculo, que comprenden, que consisten esencialmente en o que consisten en una secuencia seleccionada del grupo que consiste en: SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, un fragmento funcional de los mismos que comprende SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12 o SEQ ID NO:13, o una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias.

Los elementos reguladores de ácido nucleico para potenciar la expresión génica específica de músculo cardíaco y esquelético pueden comprender, consistir esencialmente en o consistir en una secuencia seleccionada del grupo que consiste en: SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, un fragmento funcional del mismo que comprende SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, o SEQ ID NO:6, o una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias.

Los elementos reguladores de ácido nucleico para potenciar la expresión génica específica de músculo esquelético pueden comprender, consistir esencialmente en o consistir en una secuencia seleccionada del grupo que consiste en: SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, un fragmento funcional del mismo que comprende SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12 o SEQ ID NO:13, o una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias.

También es posible preparar elementos reguladores de ácido nucleico que comprenden una secuencia artificial

combinando dos o más secuencias idénticas o diferentes dadas a conocer en el presente documento o fragmentos funcionales de las mismas. Por consiguiente, se da a conocer en el presente documento un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo que comprende al menos dos secuencias seleccionadas del grupo que consiste en: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, una secuencia que tiene al menos el 90 %, preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, o un fragmento funcional de las mismas.

También se da a conocer en el presente documento un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo, en particular la expresión génica específica de músculo cardíaco y esquelético, que comprende al menos dos secuencias seleccionadas del grupo que consiste en: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, una secuencia que tiene al menos el 90 %, preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, o un fragmento de las mismas.

También se da a conocer en el presente documento un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo, en particular la expresión génica específica de músculo esquelético, que comprende al menos dos secuencias seleccionadas del grupo que consiste en: SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, una secuencia que tiene al menos el 90 %, preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, o un fragmento de las mismas.

Por ejemplo, se da a conocer en el presente documento un elemento regulador de ácido nucleico que comprende, que consiste esencialmente en o que consiste en SEQ ID NO:1 y SEQ ID NO:5; un elemento regulador de ácido nucleico que comprende, que consiste esencialmente en o que consiste en SEQ ID NO:1 y SEQ ID NO:10; un elemento regulador de ácido nucleico que comprende, que consiste esencialmente en o que consiste en SEQ ID NO:1, SEQ ID NO:5 y SEQ ID NO:10; un elemento regulador de ácido nucleico que comprende, que consiste esencialmente en o que consiste en 2, 3, 4 o 5 repeticiones, por ejemplo, repeticiones en tándem, de SEQ ID NO:1; un elemento regulador de ácido nucleico que comprende, que consiste esencialmente en o que consiste en 2, 3, 4 o 5 repeticiones, por ejemplo, repeticiones en tándem, de SEQ ID NO:5; o un elemento regulador de ácido nucleico que comprende, que consiste esencialmente en o que consiste en 2, 3, 4 o 5 repeticiones, por ejemplo, repeticiones en tándem, de SEQ ID NO:10.

Ejemplos particulares de elementos reguladores de ácido nucleico que comprenden una secuencia artificial incluyen los elementos reguladores que se obtienen mediante la reorganización de los sitios de unión a factores de transcripción (TFBS) presentes en las secuencias dadas a conocer presente en el presente documento. Dicha reorganización puede abarcar cambiar el orden de los TFBS y/o cambiar la posición de uno o más TFBS en relación con los otros TFBS y/o cambiar el número de copias de uno o más de los TFBS. Por ejemplo, también se da a conocer en el presente documento un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo, en particular la expresión génica específica de músculo cardíaco y esquelético, que comprende sitios de unión para E2A, HNH1, NF1, C/EBP, LRF, MyoD y SREBP; o para E2A, NF1, p53, C/EBP, LRF y SREBP; o para E2A, HNH1, HNF3a, HNF3b, NF1, C/EBP, LRF, MyoD y SREBP; o E2A, HNF3a, NF1, C/EBP, LRF, MyoD y SREBP; o para E2A, HNF3a, NF1, C/EBP, LRF, MyoD y SREBP; o para HNF4, NF1, RSRFC4, C/EBP, LRF y MyoD, o NF1, PPAR, p53, C/EBP, LRF y MyoD. Por ejemplo, también se da a conocer en el presente documento un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo, en particular la expresión génica específica de músculo esquelético, que comprende sitios de unión para E2A, NF1, SRFC, p53, C/EBP, LRF y MyoD; o para E2A, NF1, C/EBP, LRF, MyoD y SREBP; o para E2A, HNF3a, C/EBP, LRF, MyoD, SREBP y Tal1_b; o para E2A, SRF, p53, C/EBP, LRF, MyoD y SREBP; o para HNF4, NF1, RSRFC4, C/EBP, LRF y SREBP; o para E2A, HNF3a, HNF3b, NF1, SRF, C/EBP, LRF, MyoD y SREBP; o para E2A, CEBP y MyoD. En ejemplos adicionales, estos elementos reguladores de ácido nucleico comprenden al menos dos, tal como 2, 3, 4 o más copias de uno o más de los TFBS mencionados.

En caso de que el elemento regulador se proporcione como un único ácido nucleico monocatenario, por ejemplo, cuando se usa un vector de AAV monocatenario, la hebra complementaria se considera equivalente a las secuencias dadas a conocer. Por tanto, también se da a conocer en el presente documento un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo que comprende, que consiste esencialmente en o que consiste en el complemento de una secuencia descrita en el presente documento, en particular una secuencia seleccionada del grupo que consiste en: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, o un fragmento funcional de las mismas.

También se da a conocer en el presente documento un elemento regulador de ácido nucleico para potenciar la hibridación de la expresión génica específica de músculo en condiciones rigurosas con un elemento regulador de ácido nucleico descrito en el presente documento, en particular con el elemento regulador de ácido nucleico que comprende, que consiste esencialmente en o que consiste en una secuencia seleccionada del grupo que consiste en: SEQ ID

NO:1, SEQ ID NO: 2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, una secuencia que tiene al menos el 90 %, preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, un fragmento funcional de las mismas, o con su complemento. Dichos elementos reguladores de ácido nucleico no necesitan tener la misma longitud que la secuencia con la que se hibridan. Preferiblemente, el tamaño de dicho elemento regulador de ácido nucleico hibridante no difiere más del 25 % en longitud, en particular el 20 % en longitud, más en particular el 15 % en longitud, más en particular el 10 % en longitud de la secuencia con la que se hibrida.

La expresión 'hibridarse en condiciones rigurosas' se refiere a la capacidad de una molécula de ácido nucleico para hibridarse con una molécula de ácido nucleico diana en condiciones definidas de temperatura y concentración de sal. Normalmente, las condiciones de hibridación estrictas están a no de más de 25°C a 30°C (por ejemplo, 20°C, 15°C, 10°C o 5°C) por debajo de la temperatura de fusión (T_m) del dúplex nativo. Los métodos de cálculo de T_m se conocen bien en la técnica. A modo de ejemplo no limitativo, las condiciones representativas de sal y temperatura para lograr una hibridación rigurosa son: 1x SSC, SDS al 0,5 % a 65°C. La abreviatura SSC se refiere a un tampón usado en disoluciones de hibridación de ácido nucleico. Un litro de la disolución tampón de SSC de reserva 20x (veinte veces concentrada) (pH 7,0) contiene 175,3 g de cloruro de sodio y 88,2 g de citrato de sodio. Un período de tiempo representativo para lograr la hibridación es de 12 horas.

Preferiblemente, los elementos reguladores descritos en el presente documento son plenamente funcionales, aunque solo tienen una longitud limitada. Esto permite su uso en vectores o casetes de expresión de ácido nucleico sin restringir indebidamente su capacidad de carga útil. Por consiguiente, el elemento regulador dado a conocer en el presente documento puede ser un ácido nucleico de 1500 nucleótidos o menos, 1000 nucleótidos o menos, 900 nucleótidos o menos, 800 nucleótidos o menos, 700 nucleótidos o menos, más preferiblemente 600 nucleótidos o menos, tal como 550 nucleótidos o menos, 500 nucleótidos o menos, 450 nucleótidos o menos, 400 nucleótidos o menos, 350 nucleótidos o menos, o 300 nucleótidos o menos (es decir, el elemento regulador de ácido nucleico tiene una longitud máxima de 1500 nucleótidos, 1000 nucleótidos, 900 nucleótidos, 800 nucleótidos, 700 nucleótidos, preferiblemente 600 nucleótidos, tal como 550 nucleótidos, 500 nucleótidos, 450 nucleótidos, 400 nucleótidos, 350 nucleótidos o 300 nucleótidos).

Sin embargo, debe entenderse que los elementos reguladores de ácido nucleico dados a conocer conservan la actividad reguladora (es decir, con respecto a la especificidad y/o actividad de transcripción) y por tanto tienen particularmente una longitud mínima de 20 nucleótidos, 25 nucleótidos, 30 nucleótidos, 35 nucleótidos, 40 nucleótidos, 45 nucleótidos, 50 nucleótidos, 100 nucleótidos, 150 nucleótidos, 200 nucleótidos, 250 nucleótidos, 300 nucleótidos, 350 nucleótidos o 400 nucleótidos.

Por ejemplo, se da a conocer en el presente documento un elemento regulador de ácido nucleico de 1000 nucleótidos o menos, preferiblemente 600 nucleótidos o menos, tal como 550 nucleótidos o menos, 500 nucleótidos o menos, 450 nucleótidos o menos, 400 nucleótidos o menos, 350 nucleótidos o menos o 300 nucleótidos o menos, para potenciar la expresión génica específica de músculo que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO:1, SEQ ID NO: 2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, una secuencia que tiene al menos el 80 %, preferiblemente al menos el 85 %, más preferiblemente al menos el 90 %, incluso más preferiblemente al menos el 95 %, tal como el 95 %, el 96 %, el 97 %, el 98 % o el 99 % de identidad con cualquiera de estas secuencias, o un fragmento funcional de las mismas.

Los elementos reguladores de ácido nucleico dados a conocer en el presente documento pueden usarse en un casete de expresión de ácido nucleico. Por consiguiente, también se da a conocer en el presente documento el uso *in vitro* o *ex vivo* de los elementos reguladores de ácido nucleico descritos en el presente documento en un casete de expresión de ácido nucleico, así como los elementos reguladores de ácido nucleico descritos en el presente documento para su uso en medicina, en el que dichos elementos reguladores de ácido nucleico están en un casete de expresión de ácido nucleico, o un vector.

Se da a conocer en el presente documento un casete de expresión de ácido nucleico que comprende un elemento regulador de ácido nucleico tal como se describe en el presente documento, unido operativamente a un promotor. El casete de expresión de ácido nucleico puede no contener un transgén. Tal casete de expresión de ácido nucleico puede usarse para dirigir la expresión de un gen endógeno. Preferiblemente, el casete de expresión de ácido nucleico comprende un elemento regulador de ácido nucleico tal como se describe en el presente documento, unido operativamente a un promotor y un transgén.

Tal como se usa en el presente documento, el término 'casete de expresión de ácido nucleico' se refiere a moléculas de ácido nucleico que incluyen uno o más elementos de control transcripcional (tales como, pero sin limitarse a, promotores, potenciadores y/o elementos reguladores, secuencias de poliadenilación e intrones) que dirigen la expresión (trans)génica en uno o más tipos celulares, tejidos u órganos deseados. Normalmente, también contendrán un transgén, aunque también se prevé que un casete de expresión de ácido nucleico dirija la expresión de un gen endógeno en una célula en la que se inserta el casete de ácido nucleico.

El término 'unido operativamente' tal como se usa en el presente documento se refiere a la disposición de diversos elementos de moléculas de ácido nucleico en relación entre sí de manera que los elementos están funcionalmente conectados y son capaces de interactuar entre sí. Tales elementos pueden incluir, sin limitación, un promotor, un potenciador y/o un elemento regulador, una secuencia de poliadenilación, uno o más intrones y/o exones y una secuencia codificante de un gen de interés que va a expresarse (es decir, el transgén). Los elementos de secuencia de ácido nucleico, cuando están apropiadamente orientados u unido operativamente, actúan juntos modulando la actividad del otro, y en última instancia pueden afectar al nivel de expresión del transgén. Por modular quiere decirse aumentar, disminuir o mantener el nivel de actividad de un elemento particular. La posición de cada elemento en relación con otros elementos puede expresarse en cuanto al extremo 5' terminal y el extremo 3' terminal de cada elemento, y la distancia entre cualquier elemento particular puede referenciarse por el número de nucleótidos intermedios, o pares de bases, entre los elementos. Tal como entiende el experto, unido operativamente implica actividad funcional, y no está necesariamente relacionado con una unión posicional natural. De hecho, cuando se usan en casetes de expresión de ácido nucleico, los elementos reguladores normalmente estarán ubicados inmediatamente en el sentido de 5' del promotor (aunque este es generalmente el caso, definitivamente no debe interpretarse como una limitación o exclusión de posiciones dentro del casete de expresión de ácido nucleico), pero este no tiene que ser necesariamente el caso *in vivo*. Por ejemplo, una secuencia de elemento regulador que se produce de manera natural en el sentido de 3' de un gen a cuya transcripción afecta es capaz de funcionar del mismo modo cuando está ubicado en el sentido de 5' del promotor. Por tanto, el efecto regulador o potenciador del elemento regulador puede ser independiente de la posición.

El casete de expresión de ácido nucleico puede comprender un elemento regulador de ácido nucleico tal como se describe en el presente documento. Alternativamente, el casete de expresión de ácido nucleico puede comprender dos o más, tal como, por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o 10, elementos reguladores de ácido nucleico tal como se describen en el documento, es decir, se combinan modularmente para potenciar su efecto regulador (y/o potenciador). Al menos dos de los dos o más elementos reguladores de ácido nucleico pueden ser idénticos o sustancialmente idénticos, tal como los dos o más elementos reguladores pueden ser idénticos o sustancialmente idénticos. Las copias de los elementos reguladores de ácido nucleico idénticos o sustancialmente idénticos pueden proporcionarse como repeticiones en tándem en el casete de expresión de ácido nucleico. Alternativamente, al menos dos de los dos o más elementos reguladores de ácido nucleico pueden ser diferentes entre sí. El casete de expresión de ácido nucleico también puede comprender una combinación de elementos reguladores de ácido nucleico idénticos y sustancialmente idénticos y elementos reguladores de ácido nucleico no idénticos.

Por ejemplo, el casete de expresión de ácido nucleico puede comprender un elemento regulador de ácido nucleico que comprende SEQ ID NO:1, y un elemento regulador de ácido nucleico que comprende SEQ ID NO:5; o el casete de expresión de ácido nucleico puede comprender un elemento regulador de ácido nucleico que comprende SEQ ID NO:1 y un elemento regulador de ácido nucleico que comprende SEQ ID NO:10; o el casete de expresión de ácido nucleico puede comprender un elemento regulador de ácido nucleico que comprende SEQ ID NO:1, un elemento regulador de ácido nucleico que comprende SEQ ID NO:5 y un elemento regulador de ácido nucleico que comprende SEQ ID NO:1; o el elemento regulador de ácido nucleico puede comprender 2, 3, 4 o 5 elementos reguladores de ácido nucleico que comprenden SEQ ID NO:1; o el elemento regulador de ácido nucleico puede comprender 2, 3, 4 o 5 elementos reguladores de ácido nucleico que comprenden SEQ ID NO:5; o el elemento regulador de ácido nucleico puede comprender 2, 3, 4 o 5 elementos reguladores de ácido nucleico que comprenden SEQ ID NO:10.

Tal como se usa en la solicitud, el término 'promotor' se refiere a secuencias de ácido nucleico que regulan, directa o indirectamente, la transcripción de secuencias codificantes de ácido nucleico correspondientes a las que están operativamente unidas (por ejemplo, un transgén o gen endógeno). Un promotor puede funcionar solo regulando la transcripción o puede actuar en concierto con una o más de otras secuencias reguladoras (por ejemplo potenciadores o silenciadores, o elementos reguladores). En el contexto de la presente solicitud, un promotor está normalmente unido operativamente a un elemento regulador tal como se da a conocer en el presente documento para regular la transcripción de un (trans)gen. Cuando un elemento regulador tal como se describe en el presente documento está unido operativamente tanto a un promotor como a un transgén, el elemento regulador puede (1) conferir un grado significativo de expresión *in vivo* específica de músculo, en particular específica de músculo cardíaco y esquelético o específica de músculo esquelético (y/o en mioblastos, miocitos o líneas celulares derivadas del músculo, en particular mioblastos cardíacos y esqueléticos o esqueléticos, miocitos cardíacos y esqueléticos o esqueléticos, o líneas celulares derivadas de músculo cardíaco o esquelético o derivadas de músculo esquelético *in vitro*) del transgén, o (2) puede aumentar el nivel de expresión del transgén en el músculo, en particular músculo cardíaco y esquelético o músculo esquelético (y/o en mioblastos, miocitos o líneas celulares derivadas del músculo, en particular mioblastos cardíacos y esqueléticos o esqueléticos, miocitos cardíacos y esqueléticos o esqueléticos, o líneas celulares derivadas de músculo cardíaco o esquelético o de músculo esquelético *in vitro*).

El promotor puede ser homólogo (es decir, de la misma especie que el animal, en particular mamífero, que va a transfectarse con el casete de expresión de ácido nucleico) o heterólogo (es decir, de una fuente distinta de la especie del animal, en particular mamífero, que va a transfectarse con el casete de expresión). Como tal, la fuente del promotor puede ser cualquier virus, cualquier organismo unicelular procarionte o eucariota, cualquier organismo vertebrado o invertebrado, o cualquier planta, o incluso puede ser un promotor sintético (es decir, que tiene una secuencia que no

se produce de manera natural), siempre que el promotor sea funcional en combinación con los elementos reguladores descritos en el presente documento. Preferiblemente, el promotor es un promotor de mamíferos, en particular un promotor murino o humano.

5 El promotor puede ser un promotor inducible o constitutivo.

El enriquecimiento de TFBS específico de músculo en los elementos reguladores de ácido nucleico dados a conocer en el presente documento, en principio, permite que los elementos reguladores dirijan la expresión específica de músculo incluso a partir de un promotor que por sí mismo no es específico de músculo. Por tanto, los elementos reguladores dados a conocer en el presente documento pueden usarse en casetes de expresión de ácido nucleico conjuntamente con su promotor natural, así como con otro promotor. Preferiblemente, los casetes de expresión de ácido nucleico dados a conocer en el presente documento comprenden un promotor específico de músculo. Esto es para aumentar la especificidad de músculo y/o evitar la fuga de expresión en otros tejidos. Los ejemplos no limitativos de promotores específicos de músculo incluyen el promotor de desmina (DES), el promotor de alfa 2 actinina (ACTN2), el promotor de filamina-C (FLNC), el promotor de ATPasa 1 de calcio del retículo sarcoplásmico/endoplásmico (ATP2A1), el promotor de troponina I de tipo 1 (TNNI1), el promotor de miosina-1 (MYH1) de cadena ligera de miosina de músculo esquelético rápido, fosforilable (MYLPF), el promotor de tropomiosina de cadena alfa 3 (TPM3), el promotor de proteína 2 que contiene dominios de repetición de anquirina (ANKRD2), el promotor de cadena pesada de la miosina (MHC), el promotor de cadena ligera de miosina (MLC), el promotor creatinina cinasa muscular (MCK), promotores musculares sintéticos descritos en Li *et al.* (1999. Nat Biotechnol. 17:241-245), tales como el promotor de SPc5-12, el promotor de creatina cinasa muscular (MCK), el promotor de dMCK y el promotor de tMCK que consisten respectivamente en un tándem doble o triple del potenciador de MCK con el promotor basal de MCK tal como se describe en Wang *et al.* (2008. Gene Ther. 15:1489-1499).

25 Más preferiblemente, el promotor es un promotor específico de músculo de mamíferos, en particular un promotor específico de músculo humano o murino.

Por ejemplo, el promotor puede ser del gen de desmina, en particular el gen de desmina murino o humano, tal como el promotor definido en SEQ ID NO: 16. Por ejemplo, el promotor de desmina murina está disponible comercialmente como pDRIVE-mDESmin (Invivogen). El promotor de desmina se expresa tanto en el músculo cardíaco como en el músculo esquelético.

El promotor puede ser un promotor específico de músculo esquelético, en particular un promotor de creatina cinasa muscular (MCK), más particularmente el promotor doble de MCK o el promotor triple de MCK que consiste en un tándem doble o triple de potenciador de MCK y promotor basal de MCK tal como se describe en Wang *et al.* (2008. Gene Ther. 15:1489-1499).

Además, no es necesario que el promotor sea el promotor del transgén en el casete de expresión de ácido nucleico, aunque es posible que el transgén se transcriba a partir de su propio promotor.

Para minimizar la longitud del casete de expresión de ácido nucleico, los elementos reguladores pueden estar unidos a promotores mínimos, o versiones acortadas de los promotores descritos en el presente documento. Un promotor mínimo (también denominado promotor basal o promotor central) tal como se usa en el presente documento forma parte de un promotor de tamaño completo que todavía es capaz de dirigir la expresión, pero que carece de al menos parte de la secuencia cuya expresión contribuye a regular (por ejemplo, específica de tejido). Esta definición abarca tanto promotores de los que se han delecionado elementos reguladores (específicos de tejido) que son capaces de dirigir la expresión de un gen pero que han perdido su capacidad para expresar ese gen de una manera específica de tejido y promotores de los que se han delecionado elementos reguladores (específicos de tejido) que son capaces de dirigir (posiblemente disminuida) la expresión de un gen, pero no que no han perdido necesariamente su capacidad de expresar ese gen de una manera específica de tejido. Preferiblemente, el promotor contenido en el casete de expresión de ácido nucleico dado a conocer en el presente documento tiene 1000 nucleótidos o menos de longitud, 900 nucleótidos o menos, 800 nucleótidos o menos, 700 nucleótidos o menos, 600 nucleótidos o menos, 500 nucleótidos o menos, 400 nucleótidos o menos, 300 nucleótidos o menos, o 250 nucleótidos o menos.

El término 'transgén' usado en el presente documento se refiere a secuencias de ácido nucleico particulares que codifican para un polipéptido o una porción de un polipéptido que va a expresarse en una célula en la que se introduce la secuencia de ácido nucleico. Sin embargo, también es posible que los transgenes se expresen como ARN, normalmente para controlar (por ejemplo, disminuir) la cantidad de un polipéptido particular en una célula en la que se inserta la secuencia de ácido nucleico. Estas moléculas de ARN incluyen, pero no se limitan a, moléculas que ejercen su función a través de interferencia de ARN (ARNhp, iARN), regulación por micro-ARN (miR) (que puede usarse para controlar la expresión de genes específicos), ARN catalítico, ARN antisentido, aptámeros de ARN, etc. La forma en que la secuencia de ácido nucleico se introduce en una célula no es esencial para la invención, puede ser, por ejemplo, a través de integración en el genoma o como plásmido episomal. Cabe destacar que la expresión del transgén puede restringirse a un subconjunto de las células en las que se introduce la secuencia de ácido nucleico. El término 'transgén' pretende incluir (1) una secuencia de ácido nucleico que no se encuentra de manera natural en la célula (es decir, una secuencia de ácido nucleico heteróloga); (2) una secuencia de ácido nucleico que es una forma mutante de

una secuencia de ácido nucleico que se encuentra de manera natural en la célula en la que se ha introducido; (3) una secuencia de ácido nucleico que sirve para añadir copias adicionales de la misma secuencia de ácido nucleico (es decir, homóloga) o una similar que se produce de manera natural en la célula en la que se ha introducido; o (4) una secuencia de ácido nucleico homóloga o que se produce de manera natural silenciosa cuya expresión se induce en la célula en la que se ha introducido.

El transgén puede ser homólogo o heterólogo para el promotor (y/o para el animal, en particular mamífero, en el que se introduce, por ejemplo, en los casos en que el casete de expresión de ácido nucleico se usa para terapia génica).

El transgén puede ser un ADNc de longitud completa o secuencia genómica de ADN, o cualquier fragmento, subunidad o mutante del mismo que tenga al menos alguna actividad biológica. En particular, el transgén puede ser un minigen, es decir, una secuencia génica que carece de parte, la mayoría o la totalidad de sus secuencias intrónicas. El transgén puede contener por tanto opcionalmente secuencias de intrón. Opcionalmente, el transgén puede ser una secuencia híbrida de ácido nucleico, es decir, una construida a partir de fragmentos de ADN genómico y/o ADNc homólogo y/o heterólogo. Por 'forma mutante' se entiende una secuencia de ácido nucleico que contiene uno o más nucleótidos que son diferentes de la secuencia silvestre o que se produce de manera natural, es decir, la secuencia de ácido nucleico mutante contiene una o más sustituciones, deleciones y/o inserciones de nucleótidos. La sustitución, delección y/o inserción de nucleótidos puede dar lugar a un producto génico (es decir, proteína o ácido nucleico) que es diferente en su secuencia de aminoácidos/ácido nucleico de la secuencia de aminoácidos/ácido nucleico silvestre. La preparación de tales mutantes se conoce bien en la técnica. En algunos casos, el transgén también puede incluir una secuencia que codifica para un péptido líder o secuencia señal de manera que el producto transgénico se secretará de la célula.

El transgén que puede estar contenido en los casetes de expresión de ácido nucleico descritos en el presente documento codifica normalmente para un producto génico tal como ARN o un polipéptido (proteína).

El transgén puede codificar para una proteína terapéutica. La proteína terapéutica puede ser una proteína secretable. Los ejemplos no limitantes de proteínas secretables, en particular proteínas terapéuticas secretables, incluyen factores de coagulación, tales como factor VIII o factor IX, insulina, eritropoyetina, lipoproteína lipasa, anticuerpos o nanocuerpos, factores de crecimiento, citocinas, quimiocinas, factores plasmáticos, etc. La proteína terapéutica también puede ser una proteína estructural. Los ejemplos no limitativos de proteínas estructurales, en particular proteínas terapéuticas estructurales, incluyen distrofina y sarcoglicanos. Por ejemplo, el transgén puede comprender el gen de microdistrofina 1 (*MD1*) o el gen de folistatina (*FST*).

El transgén puede codificar para una proteína inmunogénica. Los ejemplos no limitativos de proteínas inmunogénicas incluyen epítomos y antígenos derivados de un patógeno.

Tal como se usa en el presente documento, el término "inmunogénico" se refiere a una sustancia o composición capaz de provocar una respuesta inmunitaria.

También pueden incorporarse otras secuencias en el casete de expresión de ácido nucleico dado a conocer en el presente documento, normalmente para aumentar o estabilizar adicionalmente la expresión del producto transgénico (por ejemplo, intrones y/o secuencias de poliadenilación).

Cualquier intrón puede usarse en los casetes de expresión descritos en el presente documento. El término "intrón" abarca cualquier porción de un intrón completo que sea lo suficientemente grande como para reconocerse y cortarse y empalmarse por el aparato de corte y empalme nuclear. Normalmente, se prefieren secuencias de intrón cortas, funcionales, con el fin de mantener el tamaño del casete de expresión lo más pequeño posible, lo que facilita la construcción y manipulación del casete de expresión. El intrón puede obtenerse de un gen que codifica para la proteína que está codificada por la secuencia codificante dentro del casete de expresión. El intrón puede estar ubicado en 5' con respecto a la secuencia codificante, en 3' con respecto a la secuencia codificante, o dentro de la secuencia codificante. Una ventaja de ubicar el intrón en 5' en la secuencia codificante es minimizar la posibilidad de que el intrón interfiera con la función de la señal de poliadenilación. El casete de expresión de ácido nucleico dado a conocer en el presente documento puede comprender además un intrón. Ejemplos no limitativos de intrones adecuados son el intrón del Virus Diminuto del Ratón (MVM), el intrón de beta-globina (betaIVS-II), el intrón A del factor IX (FIX), el intrón t pequeño de virus de simios 40 (SV40) y el intrón de beta-actina.

Preferiblemente, el intrón es intrón de MVM.

Cualquier señal de poliadenilación que dirija la síntesis de una cola de poliA es útil en los casetes de expresión descritos en el presente documento, ejemplos de los cuales los conoce bien un experto en la técnica. Las señales de poliadenilación a modo de ejemplo incluyen, pero no se limitan a, secuencias de poliA derivadas del gen tardío del virus de simios 40 (SV40), la señal de poliadenilación de la hormona del crecimiento bovina (BGH), el gen mínimo de β -globina de conejo (mRBG) y el sitio de poliA sintética s(SPA) descrito en Levitt *et al.* (1989, Genes Dev 3:1019-1025) (SEQ ID NO:46).

Preferiblemente, la señal de poliadenilación se deriva de SV40 (es decir, pA de SV40).

Por ejemplo, se da a conocer un casete de expresión de ácido nucleico en el presente documento que comprende un elemento regulador de ácido nucleico que consiste en SEQ ID NO:10 o una secuencia que tiene el 95 % de identidad con dicha secuencia, unido operativamente a un promotor, preferiblemente un promotor seleccionado del grupo que consiste en el promotor del gen de desmina o el promotor de SPc5-12, y un transgén, preferiblemente un transgén que codifica para una luciferasa. El casete de expresión de ácido nucleico puede comprender además un intrón de MVM. El casete de expresión de ácido nucleico puede comprender aun adicionalmente una señal de poliadenilación, preferiblemente una señal de poliadenilación derivada de SV40.

En otro ejemplo, se da a conocer un casete de expresión de ácido nucleico que comprende un elemento regulador de ácido nucleico que consiste en SEQ ID NO:10 o una secuencia que tiene el 95 % de identidad con dicha secuencia, unido operativamente a un promotor, preferiblemente el promotor del gen de desmina, y un transgén, preferiblemente un transgén que codifica para microdistrofina 1 o folistatina. El casete de expresión de ácido nucleico puede comprender además un intrón de MVM. El casete de expresión de ácido nucleico puede comprender aun adicionalmente una señal de poliadenilación, preferiblemente en la que la señal de poliadenilación tiene SEQ ID NO:46.

El elemento regulador de ácido nucleico y el casete de expresión de ácido nucleico dados a conocer en el presente documento puede usarse como tales, o normalmente, pueden formar parte de un vector de ácido nucleico. Por consiguiente, también se da a conocer en el presente documento el uso de un elemento regulador de ácido nucleico tal como se describe en el presente documento o un casete de expresión de ácido nucleico tal como se describe en el presente documento en un vector, en particular un vector de ácido nucleico.

También se da a conocer en el presente documento un vector que comprende un elemento regulador de ácido nucleico tal como se da a conocer en el presente documento. El vector también puede comprender un casete de expresión de ácido nucleico tal como se da a conocer en el presente documento.

El término 'vector' tal como se usa en la solicitud se refiere a moléculas de ácido nucleico, por ejemplo, ADN bicatenario, que puede haberse insertado en otra molécula de ácido nucleico (la molécula de ácido nucleico de inserción), tal como, pero sin limitarse a, una molécula de ADNc. El vector se usa para transportar la molécula de ácido nucleico de inserción a una célula huésped adecuada. Un vector puede contener los elementos necesarios que permiten transcribir la molécula de ácido nucleico de inserción y, opcionalmente, convertir el transcrito en un polipéptido. La molécula de ácido nucleico de inserción puede derivarse de la célula huésped, o puede derivarse de una célula u organismo diferente. Una vez en la célula huésped, el vector puede replicarse independientemente de, o en coincidencia con, el ADN cromosómico huésped, y pueden generarse varias copias del vector y su molécula de ácido nucleico de inserción. Los vectores pueden ser vectores episomales (es decir, que no se integran en el genoma de una célula huésped), o pueden ser vectores que se integran en el genoma de la célula huésped. El término 'vector' también puede definirse por tanto como un vehículo de inserción génica que facilita la transferencia de genes a una célula diana. Esta definición incluye tanto vectores virales como no virales. Los vectores no virales incluyen, pero no se limitan a, lípidos catiónicos, liposomas, nanopartículas, PEG, PEI, vectores de plásmido (por ejemplo, vectores pUC, vectores bluescript (pBS) y pBR322 o derivados de los mismos que carecen de secuencias bacterianas (minicírculos)), vectores basados en transposones (por ejemplo, vectores PiggyBac (PB) o vectores Sleeping Beauty (SB)), etc. Los vectores virales se derivan de virus e incluyen, pero no se limitan a, vectores retrovirales, lentivirales, adenovirales, de virus adenoasociados, de virus del herpes, de virus de la hepatitis o similares. Normalmente, pero no necesariamente, los vectores virales son de replicación deficiente ya que han perdido la capacidad de propagarse en una célula dada puesto que se han eliminado genes virales esenciales para la replicación del vector viral. Sin embargo, algunos vectores virales también pueden adaptarse para replicarse específicamente en una célula dada, tal como por ejemplo una célula cancerosa, y se usan normalmente para desencadenar la (oncó)lisis específica de células (cancerosas). Los virosomas son un ejemplo no limitativo de un vector que comprende tanto elementos virales como no virales, en particular combinan liposomas con un virus VIH o influenza inactivado (Yamada *et al.*, 2003). Otro ejemplo abarca vectores virales mezclados con lípidos catiónicos.

Preferiblemente, el vector es un vector viral, tal como un vector retroviral, lentiviral, adenoviral o de virus adenoasociado (AAV), más preferiblemente un vector de AAV. Se usan preferiblemente vectores de AAV como vectores de AAV autocomplementarios, bicatenarios (scAAV) con el fin de superar una de las etapas limitantes de la transducción de AAV (es decir, conversión AAV monocatenario en bicatenario) (McCarty, 2001, 2003; Nathwani *et al.*, 2002, 2006, 2011; Wu *et al.*, 2008), aunque el uso de vectores de AAV monocatenarios (ssAAV) también se incluye en el presente documento.

El serotipo de AAV 9 (AAV9) es ideal para lograr una transducción eficaz en el corazón y el músculo esquelético. Por consiguiente, un vector particularmente preferido es un vector de AAV9, más particularmente un vector de AAV9 autocomplementario (scAAV9).

El vector también puede ser un vector no viral, preferiblemente un plásmido, un minicírculo o un vector basado en transposones, tal como un vector basado en Sleeping Beauty (SB) o un vector basado en PiggyBac (PB).

Aun alternativamente, el vector puede comprender elementos virales y no virales.

Por ejemplo, se da a conocer un vector en el presente documento que comprende un casete de expresión de ácido nucleico que comprende un elemento regulador de ácido nucleico que consiste en SEQ ID NO:10, un promotor, preferiblemente el promotor del gen de desmina, un intrón de MVM, un transgén, preferiblemente un transgén que codifica para microdistrofina 1, y una señal de poliadenilación, preferiblemente la señal de poliadenilación que tiene SEQ ID NO:46, tal como el vector que tiene SEQ ID NO: 44.

Por ejemplo, se da a conocer un vector en el presente documento que comprende un casete de expresión de ácido nucleico que comprende un elemento regulador de ácido nucleico que consiste en SEQ ID NO:10, un promotor, preferiblemente el promotor del gen de desmina, un intrón de MVM, un transgén, preferiblemente un transgén que codifica para folistatina, y una señal de poliadenilación, preferiblemente la señal de poliadenilación que tiene SEQ ID NO:46, tal como el vector que tiene SEQ ID NO: 45.

Los vectores y casetes de expresión de ácido nucleico dados a conocer en el presente documento pueden usarse, por ejemplo, para expresar proteínas que normalmente se expresan y se utilizan en el músculo (es decir, proteínas estructurales), o para expresar proteínas que se expresan en músculo y que luego se exportan al torrente sanguíneo para su transporte a otras partes del cuerpo (es decir, proteínas secretables). Por ejemplo, los vectores y casetes de expresión dados a conocer en el presente documento pueden usarse para expresar una cantidad terapéutica de un producto génico (tal como un polipéptido, en particular una proteína terapéutica, o ARN) con fines terapéuticos, en particular para terapia génica. Normalmente, el producto génico está codificado por el transgén dentro del vector o casete de expresión, aunque en principio también es posible aumentar la expresión de un gen endógeno con fines terapéuticos. En un ejemplo alternativo, los vectores y casetes de expresión dados a conocer en el presente documento pueden usarse para expresar una cantidad inmunológica de un producto génico (tal como un polipéptido, en particular una proteína inmunogénica, o ARN) con fines de vacunación.

Los vectores y casetes de expresión de ácido nucleico tal como se enseñan en el presente documento pueden formularse en una composición farmacéutica con un excipiente farmacéuticamente aceptable, es decir, una o más sustancias portadoras farmacéuticamente aceptables, y/o aditivos, por ejemplo, tampones, portadores, excipientes, estabilizadores, etc. La composición farmacéutica puede proporcionarse en forma de un kit.

El término “farmacéuticamente aceptable” tal como se usa en el presente documento es coherente con la técnica y significa compatible con otros componentes de la composición farmacéutica y no perjudicial para el receptor de la misma.

Por consiguiente, también se da a conocer en el presente documento una composición farmacéutica que comprende un casete de expresión de ácido nucleico o un vector descrito en el presente documento.

El uso de elementos reguladores del ácido nucleico descritos en el presente documento para la fabricación de estas composiciones farmacéuticas también se da a conocer en el presente documento.

La composición farmacéutica puede ser una vacuna. La vacuna puede comprender además uno o más adyuvantes para mejorar la respuesta inmunitaria. Adyuvantes adecuados incluyen, por ejemplo, pero sin limitación, saponina, geles minerales tales como hidróxido de aluminio, sustancias tensioactivas tales como lisolecitina, polioles plurónicos, polianiones, péptidos, emulsiones de aceite o hidrocarburos, bacilos Calmette-Guerin (BCG), *Corynebacterium parvum* y el adyuvante sintético QS-21. Opcionalmente, la vacuna puede comprender además una o más moléculas inmunoestimuladoras. Los ejemplos no limitativos de moléculas inmunoestimuladoras incluyen diversas citocinas, linfocinas y quimiocinas con actividades inmunoestimuladoras, inmunopotenciadoras y proinflamatorias, tales como interleucinas (por ejemplo, IL-1, IL-2, IL-3, IL-4, IL-12, IL-13); factores de crecimiento (por ejemplo, factor estimulante (CSF) de colonias de granulocitos-macrófagos (GM)); y otras moléculas inmunoestimuladoras, tales como el factor inflamatorio de macrófagos, ligando Flt3, B7.1; B7.2, etc.

Un aspecto adicional se refiere a los elementos reguladores de ácido nucleico, los casetes de expresión de ácido nucleico, los vectores o las composiciones farmacéuticas descritas en el presente documento para su uso en medicina.

Tal como se usan en el presente documento, los términos “tratar” o “tratamiento” se refieren tanto a tratamiento terapéutico como a medidas profilácticas o preventivas. Los resultados clínicos beneficiosos o deseados incluyen, pero no se limitan a, la prevención de un trastorno o estado clínico no deseado, la reducción de la incidencia de un trastorno, el alivio de síntomas asociados con un trastorno, la disminución de la extensión de un trastorno, estado estabilizado (es decir, sin empeoramiento) de un trastorno, retraso o ralentización de la progresión de un trastorno, mejora o paliación del estado de un trastorno, remisión (ya sea parcial o total), detectable o indetectable, o combinaciones de los mismos. “Tratamiento” también puede significar prolongar la supervivencia en comparación con la supervivencia esperada si no se recibe tratamiento.

Tal como se usan en el presente documento, los términos “tratamiento terapéutico” o “terapia” y similares, se refieren

a tratamientos en los que el objeto es llevar al cuerpo de un sujeto o un elemento del mismo de un trastorno o cambio fisiológico no deseado a un estado deseado, como un estado menos grave o desagradable (por ejemplo, mejora o paliación), o de vuelta a su estado normal y saludable (por ejemplo, restaurar la salud, la integridad física y el bienestar físico de un sujeto), mantenerlo en dicho cambio o trastorno fisiológico no deseado (por ejemplo, estabilización, o sin empeoramiento), o prevenir o ralentizar la progresión hasta un estado más grave o peor en comparación con dicho trastorno o cambio fisiológico no deseado.

Tal como se usa en el presente documento, los términos “prevención”, “tratamiento preventivo” o “tratamiento profiláctico” y similares abarcan la prevención de la aparición de una enfermedad o trastorno, incluyendo la reducción de la gravedad de una enfermedad o trastorno o síntomas asociados con los mismos antes de la afección por dicha enfermedad o trastorno. Tal prevención o reducción antes de la afección se refiere a la administración de los elementos reguladores de ácido nucleico, los casetes de expresión de ácido nucleico, los vectores o las composiciones farmacéuticas descritas en el presente documento a un paciente que en el momento de la administración no está afectado por síntomas de la enfermedad o trastorno. “Prevenir” comprende también la prevención de la reaparición o la prevención de recaídas de una enfermedad o trastorno, por ejemplo después de un período de mejoría. Los elementos reguladores de ácido nucleico, los casetes de expresión de ácido nucleico, los vectores o las composiciones farmacéuticas descritas en el presente documento pueden usarse en terapia génica, en particular terapia génica dirigida a músculos, más particularmente terapia génica dirigida a músculo cardíaco y esquelético, o terapia génica dirigida a músculo esquelético.

También se da a conocer en el presente documento el uso de los elementos reguladores de ácido nucleico, los casetes de expresión de ácido nucleico, los vectores o las composiciones farmacéuticas descritas en el documento para la fabricación de un medicamento para terapia génica, en particular terapia génica dirigida del músculo, más particularmente terapia génica dirigida a músculo cardíaco y esquelético o terapia génica dirigida a músculo esquelético.

También se da a conocer en el presente documento un método para la terapia génica, en particular terapia génica dirigida a músculo, más particularmente terapia génica dirigida a músculo cardíaco y esquelético o terapia génica dirigida a músculo esquelético, en un sujeto que necesita dicha terapia génica que comprende:

- introducir en el sujeto, en particular en el músculo del sujeto, más particularmente en el músculo cardíaco o esquelético del sujeto, un casete de expresión de ácido nucleico, un vector o una composición farmacéutica descrita en el presente documento, en el que el casete de expresión de ácido nucleico, el vector o la composición farmacéutica comprende un elemento regulador de ácido nucleico descrito en el presente documento unido operativamente a un promotor y un transgén; y

- expresar una cantidad terapéuticamente eficaz del producto transgénico en el sujeto, en particular en el músculo del sujeto, más particularmente en el músculo cardíaco y esquelético o en el músculo esquelético del sujeto.

El producto transgénico puede ser un polipéptido, en particular una proteína estructural tal como, por ejemplo, distrofina o un sarcoglicano, o una proteína secretable tal como, por ejemplo, un factor de coagulación, por ejemplo, factor IX o VIII, una citocina, un factor de crecimiento, un anticuerpo o nanocuerpo, una quimiocina, un factor plasmático, insulina, eritropoyetina, lipoproteína lipasa. Por ejemplo, el producto transgénico puede ser folistatina o microdistrofina, en particular microdistrofina 1. Alternativamente, el producto transgénico puede ser ARN, tal como ARNip.

Las enfermedades y trastornos a modo de ejemplo que pueden beneficiarse de la terapia génica usando los elementos reguladores de ácido nucleico, los casetes de expresión de ácido nucleico, los vectores o las composiciones farmacéuticas descritas en el presente documento incluyen distrofia muscular (por ejemplo, la distrofia muscular de Duchenne (DMD)/distrofia muscular de Becker (DMO)), distrofia miotónica, enfermedad neuromuscular, enfermedades de neuronas motoras (MND), tales como enfermedad Charcot-Marie-Tooth (CMT), atrofia muscular espinal (SMA), esclerosis lateral amiotrófica (ELA), distrofia muscular de Emery-Dreifuss, distrofia muscular facioescapulohumeral (FSHD), distrofias musculares congénitas, miopatías congénitas, distrofia muscular de la cintura y las extremidades, miopatías metabólicas, enfermedades inflamatorias musculares, miastenia, miopatías mitocondriales, anomalías de los canales iónicos, enfermedades de la envuelta nuclear, cardiomiopatías, hipertrofia cardíaca, insuficiencia cardíaca, miopatías distales, enfermedades cardiovasculares, hemofilia, incluyendo hemofilia A y B, y diabetes.

En la técnica se han descrito extensamente protocolos de terapia génica. Estos incluyen, pero no se limitan a, inyección intramuscular de plásmido (desnudo o en liposomas), inserción génica hidrodinámica en diversos tejidos, incluyendo músculo, inyección intersticial, instilación en las vías respiratorias, aplicación al endotelio, parénquima intrahepático y administración intravenosa o intraarterial. Se han desarrollado diversos dispositivos para potenciar la disponibilidad de ADN para la célula diana. Un enfoque simple consiste en poner en contacto físicamente la célula diana con catéteres o materiales implantables que contienen ADN. Otro enfoque consiste en utilizar dispositivos de inyección por chorro sin agujas que proyectan una columna de líquido directamente en el tejido diana a alta presión. Estos paradigmas de inserción también pueden usarse para suministrar vectores. Otro enfoque para la inserción génica dirigida es el uso de conjugados moleculares, que consisten en proteínas o ligandos sintéticos a los que se ha unido

un agente de unión a ácido nucleico o ADN para el direccionamiento específico de ácidos nucleicos a células (Cristiano *et al.*, 1993). Los elementos reguladores de ácido nucleico, los casetes de expresión de ácido nucleico, los vectores o las composiciones farmacéuticas descritas en el presente documento pueden usarse como vacuna, más particularmente para su uso como vacuna profiláctica.

También se da a conocer en el presente documento el uso de los elementos reguladores de ácido nucleico, los casetes de expresión de ácido nucleico, los vectores o las composiciones farmacéuticas descritas en el documento para la fabricación de una vacuna, en particular para la fabricación de una vacuna profiláctica.

También se da a conocer en el presente documento un método de vacunación, en particular vacunación profiláctica, de un sujeto que necesita dicha vacunación que comprende:

- introducir en el sujeto, en particular en el músculo del sujeto, más particularmente en el músculo cardíaco o músculo esquelético del sujeto, un casete de expresión de ácido nucleico, un vector o una composición farmacéutica descrita en el presente documento, en el que el casete de expresión de ácido nucleico, el vector o la composición farmacéutica comprende un elemento regulador de ácido nucleico descrito en el presente documento unido operativamente a un promotor y un transgén; y

- expresar una cantidad inmunológicamente eficaz del producto transgénico en el sujeto, en particular en el músculo del sujeto, más particularmente en músculo cardíaco y esquelético o en el músculo esquelético del sujeto.

Tal como se usa en el presente documento, una frase tal como “un sujeto que necesita tratamiento” incluye sujetos que se beneficiarían del tratamiento de una enfermedad o trastorno mencionado. Tales sujetos pueden incluir, sin limitación, aquellos a los que se les ha diagnosticado dicha enfermedad o trastorno, aquellos propensos a contraer o desarrollar dicha enfermedad o trastorno y/o aquellos en los que dicha enfermedad o trastorno debe prevenirse.

Los términos “sujeto” y “paciente” se usan indistintamente en el presente documento y se refieren a animales, vertebrados preferiblemente, mamíferos más preferiblemente e incluyen específicamente pacientes humanos y mamíferos no humanos. Los sujetos “mamíferos” incluyen, pero no se limitan a, humanos, animales domésticos, animales comerciales, animales de granja, animales de zoológico, animales deportivos, animales de compañía y animales de experimentación, tales como perros, gatos, cobayas, conejos, ratas, ratones, caballos, ganado, vacas; primates tales como simios, monos, orangutanes y chimpancés; cánidos tales como perros y lobos; felinos tales como gatos, leones y tigres; équidos tales como caballos, burros y cebras; animales alimentarios tales como vacas, cerdos y ovejas; ungulados tales como ciervos y jirafas; roedores tales como ratones, ratas, hámsteres y cobayas; etcétera. Los pacientes o sujetos preferidos son sujetos humanos.

Una ‘cantidad terapéutica’ o ‘cantidad terapéuticamente eficaz’ tal como se usa en el presente documento se refiere a la cantidad de producto genético eficaz para tratar una enfermedad o trastorno en un sujeto, es decir, para obtener el efecto local o sistémico deseado. El término, por tanto, se refiere a la cantidad de producto génico que provoca la respuesta biológica o medicinal en un tejido, sistema, animal o humano que está buscando un investigador, veterinario, doctor u otro médico. Tal cantidad dependerá normalmente del producto génico y la gravedad de la enfermedad, pero puede decidirla el experto, posiblemente a través de experimentación rutinaria.

Una “cantidad inmunológicamente eficaz” tal como se usa en el presente documento se refiere a la cantidad de producto (trans)génico eficaz para potenciar la respuesta inmunitaria de un sujeto contra una exposición posterior al inmunógeno codificado por el (trans)gen. Los niveles de inmunidad inducida pueden determinarse, por ejemplo, midiendo las cantidades de anticuerpos séricos y/o secretores neutralizantes, por ejemplo, mediante neutralización de placa, fijación del complemento, ensayo de inmunoabsorción ligado a enzimas o de microneutralización.

Normalmente, la cantidad de producto (trans)génico expresada cuando se usa un vector o casete de expresión tal como se describe en el presente documento (es decir, con al menos un elemento regulador de ácido nucleico específico de músculo) es mayor que cuando se usa un vector o casete de expresión idéntico pero sin un elemento regulador de ácido nucleico en el mismo. Más particularmente, la expresión es al menos el doble de alta, al menos cinco veces más alta, al menos diez veces más alta, al menos 20 veces más alta, al menos 30 veces más alta, al menos 40 veces más alta, al menos 50 veces más alta o incluso al menos 60 veces más alta que cuando se compara con el mismo vector o casete de expresión de ácido nucleico sin elemento regulador de ácido nucleico. Además, la expresión superior sigue siendo específica de músculo, en particular tanto de músculo cardíaco como esquelético o músculo esquelético solo. Además, los vectores o casetes de expresión descritos en el presente documento dirigen la expresión de una cantidad terapéutica del producto génico durante un período prolongado. Normalmente, se prevé que la expresión terapéutica dure al menos 20 días, al menos 50 días, al menos 100 días, al menos 200 días y, en algunos casos, 300 días o más. La expresión del producto génico (por ejemplo, polipéptido) puede medirse por cualquier medio reconocido en la técnica, tal como mediante ensayos basados en anticuerpos, por ejemplo, un ensayo de inmunotransferencia de tipo Western o ELISA, por ejemplo, para evaluar si se logra la expresión terapéutica del producto génico. La expresión del producto génico también puede medirse en un bioensayo que detecta una actividad enzimática o biológica del producto génico. También se da a conocer en el presente documento el uso de los elementos reguladores de ácido nucleico, los casetes de expresión de ácido nucleico o los vectores dados a conocer

en el presente documento para transfectar o transducir células musculares, preferiblemente células de músculo cardíaco y/o células de músculo esquelético.

Además se da a conocer en el presente documento el uso de los casetes de expresión de ácido nucleico o de los vectores dados a conocer en el presente documento para expresar un producto transgénico en células musculares, preferiblemente células de músculo cardíaco y/o células de músculo esquelético, en el que el ácido nucleico o el vector comprende un elemento regulador de ácido nucleico dado a conocer en el presente documento unido operativamente a un promotor y un transgén.

Además se da a conocer en el presente documento un método para expresar un producto transgénico en células musculares, preferiblemente células de músculo cardíaco y/o células de músculo esquelético, que comprende:

- transfectar o transducir las células musculares con un casete de expresión de ácido nucleico o un vector dado a conocer en el presente documento, en el que el vector o casete de expresión de ácido nucleico comprende un elemento regulador de ácido nucleico dado a conocer en el presente documento unido operativamente a un promotor y un transgén; y

- expresar el producto transgénico en las células musculares.

La transfección no viral o la transducción mediada por vectores virales de células musculares puede realizarse mediante procedimientos *in vitro*, *ex vivo* o *in vivo*. El enfoque *in vitro* requiere la transfección o transducción *in vitro* de células musculares, por ejemplo, células musculares recogidas previamente de un sujeto, líneas de células musculares o células musculares diferenciadas de, por ejemplo, células madre pluripotentes inducidas o células embrionarias no humanas. El enfoque *ex vivo* requiere la recogida de las células musculares de un sujeto, transfección o transducción *in vitro* y, opcionalmente, la reintroducción de las células musculares transfectadas en el sujeto. El enfoque *in vivo* requiere la administración del casete de expresión de ácido nucleico o del vector dado a conocer en el presente documento en un sujeto. Preferiblemente, la transfección de las células musculares se realiza *in vitro* o *ex vivo*.

El experto entiende que el uso de los elementos reguladores de ácido nucleico, los casetes de expresión de ácido nucleico y los vectores dados a conocer en el presente documento tiene implicaciones más allá de la terapia génica, por ejemplo, diferenciación coaxial de células madre en células miogénicas, modelos transgénicos para la sobreexpresión de proteínas en el músculo, etc.

La invención se explica además mediante los siguientes ejemplos.

Ejemplos

Ejemplo 1: Identificación de elementos reguladores específicos de músculo cardíaco y esquelético

Procedimientos experimentales

Se identificaron genes altamente expresados tanto en el corazón como en el músculo, pero que muestran solo una mínima expresión en otros tejidos, usando el método de medición de especificidad (SPM) descrito en Xiao *et al.* (2010. Bioinformatics 26:1273-1275). Como conjunto de datos se usó el U133A/GNF1H Atlas Gene (GSE1133) de los transcriptomas que codifican para proteínas humanas (Su *et al.* 2004. Proc. Natl. Acad. Sci. USA 101:6062-6067). Esto dio como resultado una corta lista de 5 genes: gen de filamina-c (*FLNC*), gen de actinina, alfa 2 (*ACTN2*), gen de cadena ligera 2 reguladora de miosina, isoforma ventricular/músculo cardíaco (*MYL2*), gen de desmina (*DES*) y gen de teletonina (*TCAP*). A continuación, se buscó el contexto genómico de estos genes para determinar regiones conservadas de especies cruzadas enriquecidas en sitios de unión a factores de transcripción (TFBS) asociadas con alta expresión en músculo y corazón. Se realizó filtrado adicional seleccionando los elementos reguladores de ácido nucleico supuestos que o bien se superponían o bien contenían regiones con características bioquímicas reproducibles asociados con la regulación de la transcripción y/o una estructura de cromatina abierta tal como se define en el proyecto ENCODE (The ENCODE Project Consortium. 2012. Nature 489:57-74).

Para los elementos reguladores específicos del corazón, se usó una base de datos de expresión génica de alta densidad de 18.927 genes únicos basándose en micromatrices que se derivaron de 158 muestras humanas normales de 19 órganos diferentes de 30 individuos diferentes (Son, S. *et al.* 2005. Database of mRNA gene expression profiles of multiple human organs. Genome Res. 15:443-450). Se usó esto para identificar un conjunto de genes más altamente expresados (es decir, "sobreexpresado") en el corazón en comparación con cualquiera de los otros tejidos. Se usó una prueba de la t bilateral para cada comparación por parejas. A la inversa, se identificó un conjunto de genes "subexpresados", correspondientes a aquellos genes que presentaban la expresión más baja en estos órganos respectivos en comparación con cualquiera de los otros tejidos. Este análisis dio como resultado un conjunto de 43 genes sobreexpresados y una colección de 37 genes específicos del corazón subexpresados. A continuación, se usaron las listas de identificadores (ID) de la secuencia de referencia (RefSeq) de estos genes específicos del corazón 'sobreexpresados' y 'subexpresados' para extraer las secuencias promotoras correspondientes en el sentido de 5' de

los sitios de inicio de la transcripción (TSS) notificados por 1000 bases (ensamblaje del genoma NCBI36/hg18), usando los datos de ubicación de inicio de transcripción almacenados en la tabla refGene de la base de datos del navegador de genomas de UCSC (<http://genome.ucsc.edu>).

Esto dio como resultado dos conjuntos de secuencias promotoras específicas del corazón correspondientes a los promotores de genes “sobrexpresados” o “subexpresados”. Con el fin de preparar un conjunto no redundante de secuencias promotoras representativas, las secuencias promotoras se filtraron usando ‘uclust’ (<http://www.drive5.com/uclust/>).

Los dos conjuntos de secuencias promotoras específicas de tejidos no redundantes correspondientes a los promotores de genes “sobrexpresados” o “subexpresados” se usaron como entrada para el método de DDM/MDS, descrito en detalle en De Bleser *et al.* (2007, *Genome Biol* 8: R83). El comportamiento diferencial observado podría explicarse por la presencia de uno o más elementos de TFBS característicos de los promotores del grupo de genes regulados por incremento o regulados por disminución. Estos elementos de TFBS ‘diferenciales’ pueden encontrarse usando el siguiente procedimiento. En primer lugar, se usó una biblioteca de matrices de peso posicional de TFBS (PWM) (TRANSFAC® 2010.3) para predecir TFBS en cada secuencia promotora. Para los promotores específicos de músculo se usó la aplicación Find Individual Motif Occurrences (fimo) con un corte de valor de P de 10^{-3} . El número de elementos de TFBS predichos por PWM por promotor se recogió en forma de matriz en la que cada fila corresponde a una secuencia promotora, mientras que las columnas correspondían a la PWM usada. Dos TFBS se consideraron correlacionados si sus columnas correspondientes en la matriz eran similares, lo que podría medirse usando una función de distancia. Con este enfoque, se construyeron matrices de distancia que resumen todas las asociaciones de TFBS para los TFBS en ambos grupos de promotores. Finalmente, mediante el cálculo de la matriz de diferencias de distancia (DDM) y la realización de escalado multidimensional en esta DDM para visualizar su contenido en dos dimensiones, pudieron distinguirse TFBS que no contribuían a la expresión génica diferencial observada ya que se mapearon cerca del origen de la gráfica de DDM-MDS de TFBS “que se desvían” que son probablemente responsables de la expresión génica diferencial observada. Puesto que el procedimiento de MDS representa gráficamente TFBS que están fuertemente asociados más cerca entre sí que los menos asociados, fue capaz de resaltar interacciones entre TFBS en los conjuntos de datos de promotores. Este procedimiento dio como resultado una lista de TFBS asociados con una alta expresión específica de tejido para promotores específicos del corazón.

Para los elementos reguladores específicos de músculo, se extrajo una lista de genes específicos de músculo de la base de datos Tissue-specific Gene Expression and Regulation (TiGER). Se usó esto para identificar un conjunto de los genes más altamente expresados (es decir, “sobrexpresados”) en el tejido muscular. A la inversa, se identificó un conjunto de genes “subexpresados”, correspondientes a aquellos genes que presentaban la expresión más baja en el músculo. A continuación, se usaron las listas de identificadores (ID) de la secuencia de referencia (RefSeq) de estos genes específicos de músculo ‘sobrexpresados’ y ‘subexpresados’ para extraer las secuencias promotoras correspondientes en el sentido de 5’ de los sitios de inicio de la transcripción (TSS) notificados por 1000 bases (ensamblaje del genoma NCBI36/hg18), usando los datos de ubicación de inicio de la transcripción almacenados en la tabla refGene de la base de datos de navegador del genoma de UCSC (<http://genome.ucsc.edu>). Esto dio como resultado dos conjuntos de secuencias promotoras específicas del músculo correspondientes a los promotores de genes “sobrexpresados” o “subexpresados”. Con el fin de preparar un conjunto no redundante de secuencias promotoras representativas, las secuencias promotoras se filtraron usando ‘uclust’ (<http://www.drive5.com/uclust/>).

Los dos conjuntos de secuencias promotoras específicas de tejidos no redundantes correspondientes a los promotores de genes “sobrexpresados” o “subexpresados” se usaron como entrada para el método de DDM/MDS, descrito en detalle en De Bleser *et al.* (2007, *Genome Biol* 8:R83). El comportamiento diferencial observado podría explicarse por la presencia de uno o más elementos de TFBS característicos de los promotores del grupo de genes regulados por incremento o regulados por disminución. Estos elementos de TFBS ‘diferenciales’ pudieron encontrarse usando el siguiente procedimiento. En primer lugar, se usó una biblioteca de matrices de peso posicional de TFBS (PWM) (TRANSFAC® 2010.3) para predecir TFBS en cada secuencia promotora. Para los promotores específicos de músculo se usó la aplicación Find Individual Motif Occurrences (fimo) con un corte de valor de P de 10^{-3} . El número de elementos de TFBS predichos por PWM por promotor se recogió en forma de matriz en la que cada fila corresponde a una secuencia promotora, mientras que las columnas correspondían a la PWM usada. Dos TFBS se consideraron correlacionados si sus columnas correspondientes en la matriz eran similares, lo que podría medirse usando una función de distancia. Con este enfoque, se construyeron matrices de distancia que resumen todas las asociaciones de TFBS para los TFBS en ambos grupos de promotores. Finalmente, mediante el cálculo de la matriz de diferencias de distancia (DDM) y la realización de escalado multidimensional en esta DDM para visualizar su contenido en dos dimensiones, pudieron distinguirse TFBS que no contribuían a la expresión génica diferencial observada ya que se mapearon cerca del origen de la gráfica de DDM-MDS de TFBS “que se desvían” que son probablemente responsables de la expresión génica diferencial observada. Puesto que el procedimiento de MDS representa gráficamente TFBS que están fuertemente asociados más cerca entre sí que los menos asociados, fue capaz de resaltar interacciones entre TFBS en los conjuntos de datos de promotores. Este procedimiento dio como resultado una lista de TFBS asociados con una alta expresión específica de tejido para promotores específicos de músculo.

Resultados

Este enfoque computacional condujo a la identificación de 6 secuencias reguladoras específicas de músculo cardíaco y esquelético, resumidas en la tabla 1.

Tabla 1: Elementos reguladores específicos de músculo cardíaco y esquelético. Pb: pares de bases.

Secuencia	Nombre	Gen regulado por secuencia	Tamaño (pb)	TFBS conservado presente
SEQ ID NO: 1	CSk-SH1	Des	381	E2A, HNH1, NF1, CEBP, LRF, MyoD, SREBP
SEQ ID NO: 2	CSk-SH2	Des	435	E2A, NF1, p53, CEBP, LRF, SREBP
SEQ ID NO: 3	CSk-SH3	ACTN2	551	E2A, HNH1, HNF3a, HNF3b, NF1, CEBP, LRF, MyoD, SREBP
SEQ ID NO: 4	CSk-SH4	ACTN2	430	E2A, HNF3a, NF1, CEBP, LRF, MyoD, SREBP
SEQ ID NO: 5	CSk-SH5	FLNC	454	HNF4, NF1, RSRFC4, CEBP, LRF, MyoD
SEQ ID NO: 6	CSk-SH6	FLNC	453	NF1, PPAR, p53, CEBP, LRF, MyoD

Ejemplo 2: Validación *in vivo* de elementos reguladores específicos de músculo cardíaco y esquelético mediante vectores de AAV

Procedimientos experimentales

Generación de los constructos de plásmido de AAV (pAAV-CSk-SH-Des-Luc2)

Se sintetizaron los elementos reguladores específicos de músculo cardíaco y esquelético (CSk-SH1 a CSk-SH6) mediante síntesis de oligonucleótidos convencional y se flanquearon con sitios de restricción Acc65I y MluI. Los diferentes CSk-SH se clonaron en el sentido de 5' del promotor de desmina (Des) que dirige la expresión de un gen indicador de luciferasa de luciérnaga (Luc2) en el contexto de una estructura principal de vector de virus adenoasociado (AAV) (designado como pAAV-Des-Luc2), representada esquemáticamente en la figura 3. Los constructos de plásmido de AAV correspondientes se designaron como pAAV-CSk-SH-Des-Luc2. Los plásmidos también contenían un intrón de Virus Diminuto del Ratón (MVM) y un sitio de poliadenilación (pA) del virus de simios 40 (SV40).

Producción y purificación de vectores de AAV (AAV9sc.CSk-SH1-6.Des.Luc2)

Se produjeron vectores de AAV9 mediante cotransfección con fosfato de calcio (Invitrogen Corp, Carlsbad, CA) de células renales embrionarias humanas 293T con el plásmido de pAAV de interés, un plásmido auxiliar adenoviral y un constructo de empaquetamiento quimérico que suministra el gen rep de AAV2 junto con el gen cap de AAV9, tal como se describe en Vandendriessche *et al.* (2007. J Thromb Haemost 5:16-24). Brevemente, dos días después de la transfección, se recogieron las células y se purificaron las partículas de vector usando métodos de centrifugación isopícnica. Las células recogidas se lisaron por ciclos sucesivos de congelación/descongelación y sonicación, se trataron con benzonasa (Novagen, Madison, WI) y ácido desoxicólico (Sigma-Aldrich, St. Louis, MO) y posteriormente se sometieron a 3 rondas sucesivas de ultracentrifugación en gradiente de densidad de cloruro de cesio (Invitrogen Corp, Carlsbad, CA). Se recogieron fracciones que contenían el vector de AAV, se concentraron en MgCl₂ 1 mM en solución salina tamponada con fosfato de Dulbecco (PBS) (Gibco, BRL) y se almacenaron a -80°C.

Se determinaron los títulos de vector (en genomas virales (vg)/ml) mediante PCR cuantitativa en tiempo real usando mezcla SYBR Green (que incluía tinte verde SYBR, Taqman polimerasa, ROX y dNTP todo en uno) y cebadores específicos de luciferasa en un sistema de PCR en tiempo real ABI 7500 (Applied Biosystem, Foster city, CA, EE.UU.). Los cebadores directo e inverso usados fueron 5'-CCCACCGTCGTATTCGTGAG-3' (SEQ ID NO: 14) y 5'-TCAGGGCGATGGTTTTGTCCC-3' (SEQ ID NO: 15), respectivamente.

Normalmente, para todos los vectores se obtuvieron títulos en el intervalo de 1,5-6,1x10¹¹ vg/ml a partir de un pequeño lote de producción de 20 placas de Petri de células productoras. Si se usó un mayor número de placas de Petri, tal como 60 placas de células productoras, se logró un título más alto normalmente en el intervalo de 10¹²-10¹³ gc/ml de partículas de AAV. Para generar las curvas patrón, se usaron números de copias conocidos (10²-10⁷) de los respectivos plásmidos de vector usados para generar los vectores de AAV correspondientes, que portaban los ADNc apropiados.

Estudios en animales

Todos los procedimientos con animales fueron aprobados por el comité institucional de ética animal de la Universidad Libre de Bruselas (VUB) (Bruselas, Bélgica). Todos los ratones se alojaron en condiciones específicas libres de patógenos; se proporcionaron alimentos y agua a voluntad.

Se inyectaron en ratones CB.17/lcrTac/Prkdcscid de dos días de edad vía intravenosa en la vena periorbital 50 µl de vectores concentrados (5x10⁹ vg/ratón) que contenían los diferentes elementos reguladores de CSk-SH (es decir,

CSk-SH1 a CSk-SH6) o el vector de control AAV9-Des-Luc2 (5×10^9 vg/ratón) tal como se resume en la tabla 2.

Tabla 2: Diseño del experimento para la inyección de elementos reguladores específicos de músculo cardíaco y esquelético. Vector	Número de ratón (n)	Dosis	Título (gc/ml)	Volumen de vector (μl)	Volumen de PBS (μl)	Volumen total (μl)
AAV9sc.Des.Luc2	3	5×10^9	$5,7 \times 10^{11}$	8,7	41,3	50
AAV9sc.CSk-SH1.Des.Luc2	4	5×10^9	$6,1 \times 10^{11}$	8,2	41,8	50
AAV9sc.CSk-SH2.Des.Luc2	2	5×10^9	$6,0 \times 10^{11}$	8,3	41,7	50
AAV9sc.CSk-SH3.Des.Luc2	5	5×10^9	$1,5 \times 10^{11}$	33,3	16,7	50
AAV9sc.CSk-SH4.Des.Luc2	3	5×10^9	$5,2 \times 10^{11}$	9,6	40,4	50
AAV9sc.CSk-SH5.Des.Luc2	4	5×10^9	$1,7 \times 10^{11}$	39,4	20,6	50
AAV9sc.CSk-SH6.Des.Luc2	4	5×10^9	$3,1 \times 10^{11}$	16,1	33,9	50

Se obtuvieron imágenes de los ratones entre 5 y 8 semanas después de la inyección una vez por semana usando un sistema de obtención de imágenes fotográficas *in vivo* Biospace (IVIS). La CCD se enfrió hasta -120°C y el campo de visión (FOV) se fijó a 25 cm de altura del estante de muestra. La cámara de dispositivo de carga acoplada (CCD) funciona convirtiendo los fotones que inciden en el píxel de la CCD en electrones en longitudes de onda de entre 400-100 nm, permitiendo la detección de luz infrarroja de imágenes visibles mediante anestesia con isoflurano al 2% y oxígeno. Se inyectó por vía intravenosa sustrato de D-luciferina, a una dosis de 150 μg/g de peso corporal. Se sacrificó a los ratones 9 semanas después de la inyección y se recogieron órganos intactos y se obtuvieron imágenes usando un sistema de obtención de imágenes fotográficas *in vivo* Biospace (IVIS).

Análisis de ARNm

Se extrajo el ARN de diferentes órganos de los ratones mediante un kit de purificación basado en membrana de sílice según las instrucciones del fabricante (Invitrogen Corp, Carlsbad, CA, EE.UU.). Posteriormente, se sometieron 50 ng del ARN total de cada muestra a transcripción inversa (RT) usando un kit de síntesis de ADNc (Invitrogen Corp, Carlsbad, CA, EE.UU.). A continuación, se amplificó una cantidad de ADNc correspondiente a 10 ng del ARN total mediante PCR cuantitativa (q) en un instrumento ABI 7700 (Applied Biosystems, Foster City, CA, EE.UU.), usando 5'-CCCACCGTCGTATTCGTGAG-3' (SEQ ID NO: 14) como cebador directo y 5'-TCAGGGCGATGGTTTGTCCC-3' (SEQ ID NO: 15) como cebador inverso (amplicón de 217 pb). Los patrones de qPCR consistían en plásmidos AAV9-CSk-SH-Des-Luc2 diluidos en serie de cantidad conocida. Los niveles de ARNm de Luc2 se normalizaron a los niveles de ARNm del gen endógeno de gliceraldehído-3-fosfato deshidrogenasa murina (mGAPDH), usando 5'-TGTGTCCGTCGTGGATCTGA-3' (SEQ ID NO:31) como cebador directo y 5'-GCCTGCTTCACCACCTTCTTGA-3' (SEQ ID NO:32) como cebador inverso (amplicón de 82 pb). Se amplificaron muestras de ARN con y sin transcriptasa inversa para excluir la amplificación del ADN. El tamaño de los fragmentos de PCR amplificados se verificó en un gel de agarosa al 1,8 %.

Eficiencia de transducción y biodistribución del vector

Se evaluó la eficiencia de transducción de los vectores virales y la biodistribución mediante la cuantificación de los números de copias de transgén de Luc2 en los diferentes órganos y tejidos tal como se describió anteriormente (Pacak, C.A., *et al.* 2008. Genet Vaccines Ther 6: 13). Brevemente, se extrajo ADN genómico de 30 mg de cada tejido según el protocolo del kit DNeasy Blood & Tissue (Qiagen, Chatsworth, CA, EE.UU.) y se sometieron 100 ng de ADN genómico de cada muestra a qPCR, usando el cebador directo específico de Luc2 5'-CCCACCGTCGTATTCGTGAG-3' (SEQ ID NO: 14) y el cebador inverso 5'-TCAGGGCGATGGTTTGTCCC-3' (SEQ ID NO: 15) (amplicón de 217 pb). Para generar la curva patrón, se usaron números de copia conocidos (10^2 - 10^7) del plásmido correspondiente pAAV-CSk-SH-Des-Luc2. Los resultados se expresaron como número de copias medio de AAV/100 ng de ADN genómico.

Resultados

Para evaluar el efecto *in vivo* de los elementos reguladores específicos de músculo esquelético y cardíaco (CSk-SH) identificados *in silico*, se generaron vectores adenoasociados que expresaban el gen *luc2* de luciferasa a partir de un promotor quimérico. Este promotor estaba compuesto por el promotor de desmina (Des) específico de músculo ligado a los elementos reguladores específicos de músculo cardíaco y esquelético CSk-SH1-6. Los vectores se inyectaron por vía intravenosa en ratones y se tomaron imágenes de todo el cuerpo de los ratones a las 5 y 6 semanas después de la inyección para examinar el nivel de expresión de luciferasa.

Todos los ratones en los que se inyectó un vector que comprendía un elemento regulador específico de músculo cardíaco y esquelético (CSk-SH1-6) mostraron aumento de la actividad luciferasa en comparación con ratones de control en los que se inyectó un vector de AAV9 correspondiente sin elemento regulador, indicando que todos los elementos reguladores sometidos a prueba aumentaban la expresión de luciferasa (datos no mostrado). Se observó una actividad luciferasa muy robusta y potenciada en ratones en los que inyectó el vector de AAV9 que comprendía

el elemento regulador CSk-SH1, CSk-SH3 o CSk-SH5 a las 5 y 6 semanas después de la inyección en comparación con la actividad luciferasa de los ratones de control.

También se determinó la actividad luciferasa en los órganos individuales. Nueve semanas después de la inyección, se sacrificó a los ratones y se analizaron órganos individuales para evaluar la expresión de luciferasa y el patrón de biodistribución. Se observó un aumento significativo de los niveles de actividad luciferasa en el corazón, el bíceps, el tríceps, el cuádriceps, el tibial anterior y músculo gastrocnemio de ratones en los que se inyectó un vector de AAV que comprende un elemento regulador CSk-SH1-6 en comparación con ratones de control. Se midió un aumento de hasta 48 veces de la actividad luciferasa en el corazón de ratones en los que se inyectó un vector de AAV que comprende el elemento regulador CSk-SH5 en comparación con ratones de control (figura 4A). Este elemento regulador fue también el más robusto en músculo, en particular el músculo gastrocnemio, tibial anterior, cuádriceps, bíceps y tríceps. Por tanto, el elemento regulador más potente fue CSk-SH5. El segundo elemento regulador más robusto fue CSk-SH1. Se midió un aumento de 25 veces de la actividad luciferasa en el corazón de ratones en los que se inyectó un vector de AAV que comprende CSk-SH1 en comparación con ratones de control. También se midió una actividad luciferasa robusta en otros músculos de ratones en los que se inyectó CSk-SH1. El tercer elemento regulador robusto fue CSk-SH3, seguido de CSk-SH6, CSk-SH4 y CSk-SH2. Además, la expresión fue específica para músculo cardíaco y esquelético, ya que la expresión de luciferasa estuvo ausente o limitada en todos los demás órganos (figura 5A), a pesar de la transducción del vector en los otros órganos (figura 6A-B). El nivel de actividad luciferasa medido en los órganos fue consecuente con los niveles de ARNm de luciferasa medidos en los órganos respectivos (figuras 4B, 5B). Estos datos *in vivo* muestran que los elementos reguladores de ácido nucleico CSk-SH, en particular CSk-SH1, CSk-SH3 y CSk-SH5, en particular CSk-SH1 y CSk-SH5, pueden potenciar la expresión de luciferasa específica de músculo cardíaco y esquelético.

Ejemplo 3: Identificación de elementos reguladores específicos de músculo esquelético

Procedimientos experimentales

En primer lugar, se extrajo una lista de genes específicos de músculo de la base de datos Tissue-specific Gene Expression and Regulation (TiGER). Se usó esto para identificar un conjunto de los genes más altamente expresados (es decir, "sobrexpresados") en el tejido muscular. A la inversa, se identificó un conjunto de genes "subexpresados", correspondientes a aquellos genes que presentaban la expresión más baja en el músculo. A continuación, se usaron las listas de identificadores (ID) de la secuencia de referencia (RefSeq) de estos genes específicos de músculo 'sobrexpresados' y 'subexpresados' para extraer las secuencias promotoras correspondientes en el sentido de 5' de los sitios de inicio de la transcripción (TSS) notificados por 1000 bases (ensamblaje del genoma NCBI36/hg18), usando los datos de ubicación de inicio de transcripción almacenados en la tabla refGene de la base de datos del navegador de genomas de UCSC (<http://genome.ucsc.edu>). Esto dio como resultado dos conjuntos de secuencias promotoras específicas de músculo correspondientes a los promotores de genes "sobrexpresados" o "subexpresados". Con el fin de preparar un conjunto no redundante de secuencias promotoras representativas, las secuencias promotoras se filtraron usando 'uclust' (<http://www.drive5.com/uclust/>).

Los dos conjuntos de secuencias promotoras específicas de tejidos no redundantes correspondientes a los promotores de genes "sobrexpresados" o "subexpresados" se usaron como entrada para el método de DDM/MDS, descrito en detalle en De Bleser *et al.* (2007, *Genome Biol* 8:R83). El comportamiento diferencial observado podría explicarse por la presencia de uno o más elementos de TFBS característicos de los promotores del grupo de genes regulados por incremento o regulados por disminución. Estos elementos de TFBS 'diferenciales' pueden encontrarse usando el siguiente procedimiento. En primer lugar, se usó una biblioteca de matrices de peso posicional de TFBS (PWM) (TRANSFAC® 2010.3) para predecir TFBS en cada secuencia promotora. Para los promotores específicos de músculo se usó la aplicación Find Individual Motif Occurrences (fimo) con un corte de valor de P de 10^{-3} . El número de elementos de TFBS predichos por PWM por promotor se recogió en forma de matriz en la que cada fila corresponde a una secuencia promotora, mientras que las columnas correspondían a la PWM usada. Dos TFBS se consideraron correlacionados si sus columnas correspondientes en la matriz eran similares, lo que podría medirse usando una función de distancia. Con este enfoque, se construyeron matrices de distancia que resumen todas las asociaciones de TFBS para los TFBS en ambos grupos de promotores. Finalmente, mediante el cálculo de la matriz de diferencias de distancia (DDM) y la realización de escalado multidimensional en esta DDM para visualizar su contenido en dos dimensiones, pudieron distinguirse TFBS que no contribuían a la expresión génica diferencial observada ya que se mapearon cerca del origen de la gráfica de DDM-MDS de TFBS "que se desvían" que son probablemente responsables de la expresión génica diferencial observada. Puesto que el procedimiento de MDS representa gráficamente TFBS que están fuertemente asociados más cerca entre sí que los menos asociados, fue capaz de resaltar interacciones entre TFBS en los conjuntos de datos de promotores. Este procedimiento dio como resultado una lista de TFBS asociados con una alta expresión específica de tejido para promotores específicos de músculo.

A continuación, se buscó el contexto genómico de los genes sobrexpresados específicos del tejido para determinar regiones conservadas de especies cruzadas para el TFBS asociado con una alta expresión génica específica del tejido. Para ese fin, las secuencias de todos los elementos de secuencia conservados en el ensamblaje del genoma NCBI36/hg18 se descargaron basándose en la información almacenada en la tabla phastConsElements44way de la base de datos del navegador del genoma de UCSC (<http://genome.ucsc.edu>). A los elementos de secuencia

conservados predichos se les asigna una puntuación de probabilidad logarítmica igual a su probabilidad logarítmica bajo el modelo conservado menos su probabilidad logarítmica bajo el modelo no conservado. Esto permite restringir la búsqueda de elementos reguladores supuestos que coinciden con los elementos de secuencia más conservados. Los elementos de secuencia conservados se exploraron para detectar TFBS asociados con alta expresión específica de tejido usando la aplicación fimo, tal como se describió anteriormente. Usando conjuntos de instrucciones de Perl desarrollados internamente, esto condujo a la identificación de elementos de secuencia altamente conservados que contenían agrupaciones de TFBS asociadas con una alta expresión específica de tejido (es decir, elementos reguladores de ácido nucleico). Se realizó filtrado adicional seleccionando los elementos reguladores de ácido nucleico supuestos que se superponían o contenían regiones con características bioquímicas reproducibles asociadas con la regulación de la transcripción y/o una estructura de cromatina abierta tal como se define en el proyecto ENCODE (The ENCODE Project Consortium. 2012. Nature 489:57-74).

Resultados

Este enfoque computacional condujo a la identificación de 6 secuencias reguladoras específicas de músculo, resumidas en la tabla 3.

Tabla 3: Elementos reguladores específicos de músculo. Pb: pares de bases.

Secuencia	Nombre	Gen regulado por secuencia	Tamaño (pb)	TFBS conservado presente
SEQ ID NO: 7	Sk-SH1	ATP2A1	495	E2A, NF1, SRFC, p53, CEBP, LRF, MyoD
SEQ ID NO: 8	Sk-SH2	TNNI1	344	E2A, NF1, CEBP, LRF, MyoD, SREBP
SEQ ID NO: 9	Sk-SH3	TNNI1	430	E2A, HNF3a, CEBP, LRF, MyoD, SREBP, Tal1_b
SEQ ID NO: 10	Sk-SH4	MYLPP	435	E2A, SRF, p53, CEBP, LRF, MyoD, SREBP
SEQ ID NO: 11	Sk-SH5	MYH1	474	HNF4, NF1, RSRFC4, CEBP, LRF, SREBP
SEQ ID NO: 12	Sk-SH6	TPM3	519	E2A, HNF3a, HNF3b, NF1, SRF, CEBP, LRF, MyoD, SREBP
SEQ ID NO: 13	Sk-SH7	ANKRD2	372	E2A, CEBP, MyoD

Ejemplo 4: Validación *in vivo* de elementos reguladores específicos de músculo esquelético a través de vectores de AAV que comprenden un promotor de desmina

Procedimientos experimentales

Se generaron constructos de plásmido de AAV que comprenden un elemento regulador específico de músculo esquelético (pAAV-Sk-SH-Des-Luc2) según el protocolo descrito en el ejemplo 2. Brevemente, se sintetizaron los elementos reguladores específicos de músculo esquelético mediante síntesis de oligonucleótidos convencional y se flanquearon con sitios de restricción Acc65I y MluI (Sk-SH1, 2, 3, 4, 5 y 7) o sitios de restricción BsiWI y MluI (Sk-SH6). Los diferentes Sk-SH se clonaron en el sentido de 5' del promotor de desmina (Des) que dirige la expresión de un gen indicador de luciferasa de luciérnaga (Luc2) en el contexto de la estructura principal de vector de AAV pAAV-Des-Luc2. Los plásmidos también contenían un intrón de Virus Diminuto del Ratón (MVM) y un sitio de poliadenilación (pA) del virus de simios 40 (SV40).

Se llevaron a cabo estudios de producción y purificación de vectores de AAV y en animales tal como se describió en el ejemplo 2. El diseño del experimento se resume en la tabla 4.

Tabla 4: Diseño del experimento para la inyección de elementos reguladores específicos de músculo esquelético.

Vector	Número de ratón (n)	Dosis	Título (gc/ml)	Volumen de vector (μl)	Volumen de PBS (μl)	Volumen total (μl)
AAV9sc. Des.Luc2	5	5x10 ⁹	5,7x10 ¹¹	8,7	41,3	50
AAV9sc. Sk-SH1.Des.Luc2	4	5x10 ⁹	1,7x10 ¹¹	29,4	20,6	50
AAV9sc. Sk-SH2.Des.Luc2	4	5x10 ⁹	1,8x10 ¹¹	27,7	22,3	50
AAV9sc. Sk-SH3.Des.Luc2	4	5x10 ⁹	1,9x10 ¹¹	26,3	23,7	50
AAV9sc. Sk-SH4.Des.Luc2	4	5x10 ⁹	1,6x10 ¹¹	31,3	18,7	50
AAV9sc. Sk-SH5.Des.Luc2	1	5x10 ⁹	8,1x10 ¹¹	62	3	65
AAV9sc. Sk-SH6.Des.Luc2	3	5x10 ⁹	1,4x10 ¹¹	35,7	14,3	50
AAV9sc. Sk-SH7.Des.Luc2	4	5x10 ⁹	1,8x10 ¹¹	27,7	22,3	50

Se evaluaron el análisis de ARNm y la eficiencia de transducción y la biodistribución del vector tal como se describió en el ejemplo 2.

Resultados

Se generaron vectores adenoasociados que expresaban el gen de luciferasa *luc2* a partir de un promotor quimérico compuesto por el promotor de desmina (Des) ligado a los elementos reguladores específicos de músculo esquelético Sk-SH1-7. Los vectores se inyectaron por vía intravenosa en ratones y se tomaron imágenes de todo el cuerpo de los ratones a las 5 y 6 semanas después de la inyección para examinar el nivel de expresión de luciferasa.

Todos los ratones en los que se inyectó un vector que comprende un elemento regulador específico de músculo esquelético (Sk-SH1-7) mostraron un aumento de la actividad luciferasa en comparación con ratones de control en los que se inyectó un vector de AAV9 correspondiente sin un elemento regulador, indicando que todos los elementos reguladores sometidos a prueba aumentaron la expresión de luciferasa a partir del promotor de desmina (datos no mostrados). Se observó una actividad luciferasa muy robusta y potenciada en ratones en los que inyectó el vector de AAV9 que comprende el elemento regulador Sk-SH1 o Sk-SH4 a las 5 y 6 semanas después de la inyección en comparación con la actividad luciferasa de los ratones de control.

7 semanas después de la inyección, se sacrificó a los ratones y se analizaron órganos individuales para evaluar la expresión de luciferasa y el patrón de biodistribución (figuras 8 y 9). Se midió un aumento de hasta 200 a 400 veces de la actividad luciferasa en músculo esquelético, en particular músculo gastrocnemio, tibial interior, cuádriceps, bíceps y tríceps, de ratones en los que se inyectó el vector de AAV que comprende el elemento regulador Sk-SH4 en comparación con ratones de control (figura 8). Todavía hubo un aumento de 36 veces de la actividad luciferasa medida en el corazón de dichos ratones (figura 8), debido al uso del promotor de desmina, que es específico tanto para músculo cardíaco como para el esquelético. La expresión de luciferasa estuvo ausente o limitada en todos los demás órganos distintos del tejido de músculo cardíaco y esquelético (figura 9). El segundo elemento regulador más robusto fue Sk-SH1, que también aumentó específicamente la expresión de luciferasa del promotor de desmina en el músculo esquelético, y en menor medida en el músculo cardíaco.

El nivel de actividad luciferasa medido en el corazón y los diferentes músculos esqueléticos fue consecuente con los niveles de ARNm de luciferasa medidos en los órganos respectivos (figura 12). A pesar de la transducción del vector en diferentes órganos, tal como se muestra en la figura 13, la expresión de luciferasa estuvo ausente o limitada en todos los demás órganos distintos del músculo cardíaco y esquelético (figura 9).

Estos datos *in vivo* muestran que los elementos reguladores de ácido nucleico Sk-SH, en particular Sk-SH1 y Sk-SH4, pueden potenciar la expresión de luciferasa específica de músculo cardíaco y esquelético. El elemento regulador del ácido nucleico Sk-SH4 es, con mucho, el elemento más robusto que condujo al nivel más alto de expresión de luciferasa en el músculo cardíaco y esquelético en comparación con los otros 5 elementos reguladores que se identificaron. La actividad más alta que se midió fue una regulación por incremento de 400 veces de la actividad luciferasa en el tibial.

Ejemplo 5: Comparación *in vivo* de elementos reguladores específicos de músculo y promotor de CMV

Procedimientos experimentales

Se generaron constructos de plásmido de AAV que comprendían un elemento regulador específico de músculo (pAAV-

Sk-SH4-Des-Luc2, pAAV-CSk-SH1-Des-Luc2, pAAV-CSk-SH5-Des-Luc2) según el protocolo descrito en el ejemplo 2. Brevemente, los elementos reguladores específicos de músculo se sintetizaron mediante síntesis de oligonucleótidos convencional y se clonaron en el sentido de 5' del promotor de desmina (Des) que dirige la expresión de un gen indicador de luciferasa de luciérnaga (Luc2) en el contexto de la estructura principal de vector de AAV pAAV-Des-Luc2. Los plásmidos también contenían un intrón de Virus Diminuto del Ratón (MVM) y un sitio de poliadenilación (pA) del virus de simios 40 (SV40).

Se generó un constructo de plásmido pAAVsc-CMV-Luc2-SV40pA, en el que el promotor del citomegalovirus (CMV) (SEQ ID NO: 30) se clona en el sentido de 5' del gen indicador de luciferasa de luciérnaga (Luc2) en lugar del promotor de desmina en el contexto de la estructura principal de vector de AAV pAAV-Des-Luc2. El plásmido también contenía un sitio de poliadenilación (pA) del virus de simios 40 (SV40). En la figura 10A se muestra una representación esquemática de pAAVsc-CMV-Luc2-SV40pA. Para la clonación del AAVsc-CMV-Luc2-SV40pA, se sintetizó el fragmento 5'-Acc65I-CMV-HindIII-3' y se clonó en un vector de AAVsc-Luc2-SV40pA que se restringió con el mismo par de enzimas (es decir, Acc65I y HindIII).

La producción y purificación de vectores de AAV se llevaron a cabo tal como se describió en el ejemplo 2. En ratones macho CB17 SC SCID de 6 semanas que tienen un peso promedio de $17,4 \pm 1,6$ g se inyectaron por vía intravenosa en la vena periorbital 1×10^{10} vg/ratón de vectores concentrados. Se obtuvieron imágenes de los ratones 4 semanas después de la inyección, tal como se describe en el ejemplo 2. Se sacrificó a los ratones 5 semanas después de la inyección y se recogieron los órganos intactos y se obtuvieron imágenes tal como se describió en el ejemplo 2.

Resultados

Se comparó actividad luciferasa en ratones en los que se inyectó un vector que expresaba el gen de luciferasa *luc2* a partir del promotor de desmina (Des) operativamente unido a un elemento regulador específico de músculo: Sk-SH4, CSk-SH1 o CSk-SH5, frente a ratones en los que se inyectó un vector que expresaba el gen de luciferasa *luc2* a partir del promotor de citomegalovirus (CMV). El promotor de CMV se considera uno de los promotores más potentes conocidos para permitir una expresión génica robusta en el corazón. No estaba presente ningún elemento regulador específico de músculo en estos últimos vectores.

Las figuras 10A y 10B muestran que los vectores que comprenden un elemento regulador específico de músculo tal como se describe en el presente documento operativamente unido al promotor de desmina permiten una expresión mucho mejor que el vector que comprende el promotor de CMV.

Ejemplo 6: Validación *in vivo* de elementos reguladores específicos de músculo esquelético mediante vectores de AAV que comprenden un promotor específico de músculo esquelético

Procedimientos experimentales

Se construyeron plásmidos de AAV tal como se describió en el ejemplo 4, en los que el promotor de desmina se reemplaza por el promotor específico de músculo descrito en Wang *et al.* (2008. Gene Ther. 15:1489-1499).

La producción y purificación de vectores de AAV y los estudios en animales se llevan a cabo tal como se describió en el ejemplo 2.

Resultados

Al usar un promotor específico de músculo esquelético en lugar del promotor de desmina, el aumento de la expresión de luciferasa se limita únicamente al músculo esquelético, mostrando que los elementos reguladores de ácido nucleico Sk-SH1-7, en particular Sk-SH4 y Sk-SH1 potencian la expresión de luciferasa específica de músculo esquelético.

Ejemplo 7: Validación *in vivo* de fragmentos funcionales de los elementos reguladores específicos de músculo identificados

Procedimientos experimentales

Los elementos reguladores de ácido nucleico identificados CSk-SH1 (SEQ ID NO: 1; 381 pb), CSk-SH5 (SEQ ID NO:5; 454 pb) y Sk-SH4 (SEQ ID NO:10; 435 pb) se dividieron en fragmentos funcionales más pequeños. Los sitios de unión a factores de transcripción (TFBS) de los elementos reguladores se mapearon en la secuencia respectiva (figuras 11A-C), y los fragmentos funcionales se generaron agrupando aleatoriamente regiones con TFBS. Se generaron los siguientes fragmentos funcionales:

Fragmentos funcionales de SEQ ID NO:1:

Agrupación 1A: la secuencia de nucleótidos desde la posición 33 hasta 58 en SEQ ID NO:1

Agrupación 1B: la secuencia de nucleótidos desde la posición 90 hasta 142 en SEQ ID NO:1

Agrupación 1C: la secuencia de nucleótidos desde la posición 143 hasta 233 en SEQ ID NO:1

5 Agrupación 1D: la secuencia de nucleótidos desde la posición 240 hasta 310 en SEQ ID NO:1

Agrupación 1E: la secuencia de nucleótidos desde la posición 90 hasta 233 en SEQ ID NO:1

Fragmentos funcionales de SEQ ID NO:5:

10

Agrupación 5A: la secuencia de nucleótidos desde la posición 47 hasta 130 en SEQ ID NO:5

Agrupación 5B: la secuencia de nucleótidos desde la posición 252 hasta 293 en SEQ ID NO:5

15 Agrupación 5C: la secuencia de nucleótidos desde la posición 330 hasta 450 en SEQ ID NO:5

Fragmentos funcionales de SEQ ID NO:10:

20

Agrupación 4A: la secuencia de nucleótidos desde la posición 10 hasta 180 en SEQ ID NO:10 (SEQ ID NO: 37);

Agrupación 4B: la secuencia de nucleótidos desde la posición 190 hasta 240 en SEQ ID NO:10 (SEQ ID NO: 38);

Agrupación 4C: la secuencia de nucleótidos desde la posición 241 hasta 300 en SEQ ID NO:10 (SEQ ID NO: 39);

25 Agrupación 4E: la secuencia de nucleótidos desde la posición 241 hasta 360 en SEQ ID NO:10 (SEQ ID NO:41)

Agrupación 4D: la secuencia de nucleótidos desde la posición 380 hasta 420 en SEQ ID NO:10 (SEQ ID NO: 40)

Los fragmentos funcionales de Sk-SH4 se resumen en la tabla 5.

30

Tabla 5: Fragmentos funcionales del elemento regulador específico de músculo Sk-SH4. Pb: pares de bases.

Secuencia	Nombre	Gen regulado por secuencia	Tamaño (pb)	TFBS conservado presente
SEQ ID NO: 37	Sk-SH4 ^a	MYLPF	171	CEBP, E2A, LRF, SRFb, SRFc
SEQ ID NO: 38	Sk-SH4 ^b	MYLPF	51	CEBP, E2A, LRF, SRFb
SEQ ID NO: 39	Sk-SH4 ^c	MYLPF	60	E2A, LRF, SRFb, MyoD
SEQ ID NO: 40	Sk-SH4 ^d	MYLPF	41	LRF, SRFb, p53, MyoD
SEQ ID NO: 41	Sk-SH4 ^e	MYLPF	120	CEBP, E2A, LRF, SRFb

35 Los fragmentos funcionales de Sk-SH4 se sintetizaron por síntesis de oligonucleótidos convencional y se clonaron en el vector pAAV9sc-Des-Luc2 tal como se describe en el ejemplo 2. La producción y purificación de vectores de AAV se llevaron a cabo tal como se describió en el ejemplo 2. En ratones adultos CB17/lcrTac/Prkdc scid de 6 semanas de edad con un peso promedio de aproximadamente 17-18 g por ratón se inyectaron por vía intravenosa los vectores de AAV a una dosis de 1×10^{10} vg por ratón. Después de 2 y 4 semanas después de la inyección, se sometió a los ratones a obtención de imágenes usando el mismo sistema de obtención de imágenes fotográficas *in vivo* Biospace tal como se describe en el ejemplo 2.

40

Resultados

45 Los resultados mostrados en la figura 14 muestran que los 5 fragmentos del elemento regulador Sk-SH4 fueron capaces de aumentar la expresión del gen de luciferasa dirigido a partir del promotor de desmina en comparación con el constructo de referencia sin ningún elemento regulador. El fragmento Sk-SH4^b mostró la expresión de luciferasa más alta en comparación con los otros fragmentos. Sin embargo, ninguno de los fragmentos pudo lograr una expresión de luciferasa mayor o igual en comparación con la expresión inducida por el fragmento Sk-SH4 de longitud completa.

50

Ejemplo 8: Validación *in vivo* de módulos de elementos reguladores específicos de músculo

Los elementos reguladores específicos de músculo se validan adicionalmente preparando diferentes combinaciones de los elementos reguladores y/o usando varias copias del mismo elemento regulador (por ejemplo, CSK-SH1 combinado con CSK-SH5; CSK-SH1 combinado con Sk-SH4; CSK-SH1 combinado con CSK-SH5 y Sk-SH4; CSK-SH1 repetido 3x o 4x; CSK-SH5 repetido 3x o 4x; Sk-SH4 repetido 3x o 4x). Estos constructos se incorporan en el sentido de 5' del promotor Des según el protocolo descrito en el ejemplo 2. La producción y purificación de vectores de AAV y los estudios en animales se llevan a cabo tal como se describió en el ejemplo 2.

55

Se preparan combinaciones adicionales con fragmentos funcionales de los elementos reguladores específicos de músculo identificados tal como se generaron en el ejemplo 7.

5 Ejemplo 9: Unión de factores de transcripción a elementos reguladores específicos de músculo

Procedimientos experimentales

10 En ratones de 4 semanas de edad se inyectaron por vía intravenosa en la vena periorbital vectores de AVV que contenían un elemento regulador Sk-SH4 y CSK-SH5, es decir, AAV9sc-Sk-SH4-Des-Luc y AAV9sc-CSK-SH5-Des-Luc, a una dosis de 5×10^9 vg/por ratón.

15 Se sumergieron tejidos cardíacos y musculares recogidos de los ratones en solución salina tamponada con fosfato (PBS) que contenía formaldehído al 1%, se cortaron en trozos pequeños y se incubaron a temperatura ambiente durante 15 minutos. La fijación se detuvo mediante la adición de glicina 0,125 M (concentración final). Entonces se trataron los trozos de tejido con un cortador de tejido y finalmente se centrifugaron y se lavaron dos veces en PBS. Se aisló la cromatina mediante la adición de tampón de lisis, seguida por disrupción con un homogeneizador Dounce. Los lisatos se sonicaron y el ADN se fragmentó hasta una longitud media de 300-500 pb. Se preparó ADN genómico (entrada) tratando alícuotas de cromatina con ARNasa, proteinasa K y calor para la desreticulación, seguido de la precipitación con etanol. Se resuspendieron los sedimentos y el ADN resultante se cuantificó en un espectrofotómetro NanoDrop. La extrapolación al volumen original de cromatina permitió la cuantificación del rendimiento total de cromatina. Se preclaró una alícuota de cromatina (30 μ g) con perlas de agarosa de proteína A (Invitrogen, número de catálogo 15918-014). Se aislaron regiones de ADN genómico de interés usando 4 μ g de anticuerpo contra MEF2 (Santa Cruz, sc-313), SRF (Santa Cruz, sc-335) y CEBP (Santa Cruz, sc-150). Los complejos se lavaron, se eluyeron de las perlas con tampón SDS y se sometieron a tratamiento con ARNasa y proteinasa K. Se revirtieron las reticulaciones mediante incubación durante la noche a 65°C, y se purificó el ADN de ChIP mediante extracción con fenol-cloroformo y precipitación con etanol. Se llevaron a cabo reacciones de PCR cuantitativa (qPCR) por triplicado en regiones genómicas específicas usando SYBR Green Supermix (Bio-Rad). Las señales resultantes se normalizaron para la eficiencia de cebadores llevando a cabo qPCR para cada par de cebadores usando el ADN de entrada. Las secuencias de cebadores para Sk-SH4 fueron: 5'-GTCCCTCACTCCCACTCAG-3' (SEQ ID NO: 33; directo) y 5'-GAGGAGAAGGAGATCAGACACTG-3' (SEQ ID NO: 34; inverso); para CSK-SH5: 5'-TAGCTGGGCCTTCTCTCTC-3' (SEQ ID NO: 35; directo) y 5'-CGTCTCCCTAGCAGCAACAG-3' (SEQ ID NO: 36; inverso). Se adquirieron cebadores de control negativo de Active Motif (Carlsbad, CA, EE.UU.) (n.º 71012) y son específicos para secuencias génicas no transcritas en el cromosoma 17.

Resultados

40 El elemento regulador específico de músculo Sk-SH4 contiene varios sitios de unión a factores de transcripción (TFBS) supuestos, incluyendo E2A, SRF, p53, CEBP, LRF, MyoD y SREBP. Usando un ensayo de inmunoprecipitación de cromatina (CHIP), se confirmó la unión de CEBP y SRF en el elemento Sk-SH4 en el músculo cardíaco y esquelético de ratones en los que se inyectó el vector de AAV9sc-Sk-SH4-Des-Luc (figuras 15A, B). De manera similar, se mostró la unión del elemento CEBP y MEF2 sobre el elemento regulador CSK-SH5 en el músculo cardíaco y esquelético a partir de ratones en los que se inyectó AAV9sc-CSK-SH5-Des-Luc (figuras 15C, D).

45 Ejemplo 10: Evaluación terapéutica de elementos reguladores específicos de músculo mediante vectores de AAV que comprenden un promotor de desmina

50 Se evaluó el efecto del elemento regulador específico de músculo Sk-SH4 sobre la expresión y eficacia terapéutica de dos genes terapéuticos, en particular microdistrofina y folistatina, que tienen potencial terapéutico para el tratamiento de enfermedades musculares tales como distrofia muscular de Duchenne (DMD) mediante terapia génica. Los ratones MDX-SCID replican las manifestaciones de la enfermedad de distrofia muscular de Duchenne en pacientes y, por tanto, son adecuados para evaluar la eficacia terapéutica de los vectores de AAVss-Sk-SH4-Des-MVM-MD1 y AAVss-Sk-SH4-Des-MVM-FST-2A-Luc.

Procedimientos experimentales

Clonación de genes de folistatina y MD1 en AAV

60 Los genes de microdistrofina 1 (MD1) (Koo *et al.*, 2011) y folistatina (FST) (Kota *et al.*, 2009) se clonaron y se dirigieron a partir del promotor de desmina, que estaba operativamente unido al elemento regulador específico de músculo Sk-SH4. El gen MD1 estaba flanqueado por los sitios de restricción MluI y XhoI en los extremos 5' y 3', mientras que el gen FST estaba flanqueado por los sitios de restricción MluI y Sall en los extremos 5' y 3', respectivamente y se sintetizaron por síntesis de oligonucleótidos convencional. El elemento regulador Sk-SH4, operativamente unido al promotor de desmina, se clonó en el sentido de 5' del intrón de MVM en el contexto de una estructura principal de vector viral adenoasociado monocatenario (AAVss). El vector también contenía un sitio de poliadenilación sintético Proudfoot de 49 pb (Levitt N *et al.*, 1989). El gen folistatina se ligó a un gen indicador de luciferasa mediante un

polipéptido 2A. Los constructos generados se designaron pAAVss-Sk-SH4-Des-MVM-MD1 y pAAVss-Sk-CRM4-Des-MVM-FST-2A-Luc, respectivamente, y se muestran esquemáticamente en la figura 16.

La producción y purificación de vectores de AAV se llevaron a cabo tal como se describió en el ejemplo 2.

Prueba de cinta de correr para la corrección fenotípica de ratones MDX-SCID

En ratones MDX-SCID de 4 semanas de edad (criados internamente) se inyectó por vía intravenosa en un volumen total de 100 μ l con una dosis de 2×10^{10} vg por ratón de los vectores de AAVss-Sk-SH4-Des-MVM-MD1 y AAVss-Sk-SH4-Des-MVM-FST-2A-Luc, de manera individual o en combinación. 8 semanas después de la inyección, los ratones MDX-SCID tratados y de control (en los que se inyectó PBS) se sometieron a la prueba de cinta de correr. Las pruebas de cinta de correr se realizaron usando la cinta de correr abierta Exer-3/6 (Columbus instruments, EE.UU.). La inclinación de la cinta se ajustó a 10° cuesta arriba antes de realizar la prueba. La velocidad inicial se fijó en 10 m/min y, después de eso, la velocidad se incrementó en 1 m cada minuto. La prueba se terminó en un punto en el que los ratones se sentaron durante ≥ 5 segundos en la parrilla de pulsos. En ese punto se registró la distancia cubierta por los ratones y la distancia total cubierta por los ratones durante el transcurso de la prueba se calculó usando la fórmula, distancia = $((N+n)/2) \times (N-n+1)$ donde N es el tiempo (en min) en el punto de terminación de la prueba y n es el tiempo (en minutos) al inicio de la prueba.

Análisis de ARNm

El análisis de ARNm se realizó tal como se describió en el ejemplo 2. Para cuantificar el ARNm de los genes de microdistrofina y folistatina, se usó un método basado en qPCR usando los cebadores 5'-GTGCCCTACTACATCAA-3' (SEQ ID NO:42) como cebador directo y 5'-AGGTTGTGCTGGTCCA-3' (SEQ ID NO:43) como cebador inverso (amplicón de 206 pb) para la microdistrofina, y el cebador directo específico de Luc2 5'-CCCACCGTCGTATTCTGTGAG-3' (SEQ ID NO: 14) y el cebador inverso 5'-TCAGGGCGATGTTTGTCCC-3' (SEQ ID NO: 15) (amplicón de 217 pb) para el gen de folistatina.

Resultados

Los resultados de la prueba de cinta de correr mostraron claramente que los ratones tratados mediante terapia génica con o bien MD1 o bien FST o bien la combinación de ambos, tuvieron un mejor rendimiento que los ratones de control MDX-SCID no tratados (figura 17). Los ratones MDX-SCID en los que se inyectaron el vector de AAVss-Sk-SH4-Des-MVM-MD1 o el AAVss-Sk-SH4-Des-MVM-FST-2A-Luc cubrieron dos veces la distancia cubierta por los ratones de control. Cuando ambos vectores se coinyectaron en los ratones MDX-SCID, se cubrió una distancia de 4 km, que es 4 veces mayor que la distancia cubierta por los ratones MDX-SCID de control en los que se inyectó PBS. Estos resultados demuestran claramente que la eficacia terapéutica puede lograrse usando estos vectores.

El efecto fisiológico significativo de la terapia génica usando el vector de AAVss-Sk-SH4-Des-MVM-MD1 o el AAVss-Sk-SH4-Des-MVM-FST-2A-Luc o su uso combinado se confirmó mediante examen histológico. Las secciones musculares teñidas con hematoxilina/eosina se puntuaron para el % de células nucleadas centralmente para los diferentes grupos experimentales (figura 18). Las secciones musculares de ratones silvestres C57B6 mostraron una localización predominantemente periférica del núcleo en las miofibras (datos no mostrado) y un porcentaje mínimo (< 2%) de los núcleos estaban ubicados centralmente (figura 18B). En ratones MDX/SCID no tratados, de control, los núcleos estaban predominantemente (alrededor del 50% de los núcleos) ubicados centralmente en las miofibras transectadas transversalmente (figura 18Aa, figura 18B). Ratones MDX/SCID en los que se inyectaron AAVss-Sk-SH4-Des-MVM-MD1 (AAV9-MD1) (figura 18Ac, figura 18B) o AAVss-Sk-SH4-Des-MVM-FST-2A-Luc (AAV9-FST) (figura 18Ab, figura 18B), mostraron todos una disminución del número de miofibras que están nucleadas centralmente. La administración conjunta de ambos vectores de AAVss-Sk-SH4-Des-MVM-MD1 y AAVss-Sk-SH4-Des-MVM-FST-2A-Luc, dio como resultado un cambio incluso más profundo en la localización nuclear hacia la periferia de las miofibras, y una reducción adicional en el porcentaje de núcleos ubicados centralmente (figura 18Ad, figura 18B). Estos resultados son indicativos de un estímulo regenerativo reducido y corrección fenotípica después de la terapia génica. En consecuencia, el cambio de la localización nuclear de una ubicación central a otra más periférica en las miofibras es consecuente con la corrección fenotípica tal como se muestra por mejoría de la fuerza muscular y la movilidad en la prueba de cinta de correr.

Usando RT-PCR cuantitativa, se sometió a ensayo la expresión de los genes de microdistrofina (MD1) y de folistatina (FST) en biopsias de tejidos cardíacos y musculares, en particular gastrocnemio y cuádriceps, de los ratones en los que se inyectó el constructo pAAVss-Sk-SH4-Des-MVM-MD1, el constructo pAAVss-Sk-SH4-Des-MVM-FST-2A-Luc o la combinación de ambos constructos (figura 19).

Ejemplo 11: Comparación *in vivo* de elementos reguladores específicos de músculo con promotores de CMV y SPc5-12

Procedimientos experimentales

La producción y purificación de vectores de AAV se llevaron a cabo tal como se describió en el ejemplo 2.

El vector de AAVsc-Des-MVM-Luc es un vector de AAV autocomplementario que contiene un transgén de luciferasa (Luc) dirigido a partir del promotor de desmina. La estructura principal del vector de scAAV también contenía un intrón del Virus Diminuto del Ratón (MVM) y un sitio de poliadenilación (pA) del virus de simios 40 (SV40).

El vector de AAVsc-SPc5-12-MVM-Luc tenía la misma configuración que el vector de AAVsc-Des-MVM-Luc, pero el promotor de desmina se reemplazó por el promotor de SPc5-12.

Se generaron vectores de AAV que comprendían el elemento regulador específico de músculo Sk-SH4 según el protocolo descrito en el ejemplo 2. Brevemente, el elemento regulador específico de músculo Sk-SH4 se sintetizó por síntesis de oligonucleótidos convencional y se clonó en el sentido de 5' del promotor de desmina o el SPc5-12 en el contexto de la estructura principal de vector de AAV AAVsc-Des-MVM-Luc o AAVsc-SPc5-12-MVM-Luc, respectivamente.

El vector de AAVsc-CMV-Luc se generó mediante clonación del promotor de citomegalovirus (CMV) (SEQ ID NO: 30) en el sentido de 5' del gen indicador de luciferasa en lugar del promotor de desmina en el contexto de la estructura principal de vector de AAV de AAVsc-Des-Luc, que también contenía un sitio de poliadenilación (pA).

En la figura 20 se muestra una representación esquemática de los diferentes vectores.

En ratones adultos CB17/lcrTac/Prkdcscid se inyectaron por vía intravenosa los diferentes vectores (n=3) a una dosis de 1×10^{10} vg/ratón.

El análisis de ARNm se realizó tal como se describió en el ejemplo 3.

Resultados

Se compararon los niveles de ARNm de luciferasa en ratones en los que se inyectó un vector que expresaba el gen de luciferasa a partir del promotor de desmina operativamente unido al elemento regulador específico de músculo Sk-SH4 frente a ratones en los que se inyectó un vector que expresaba el gen de luciferasa a partir de los promotores de citomegalovirus (CMV) o el SPc5-12, que no estaban operativamente unidos a un elemento regulador específico de músculo. Los promotores de CMV y SPc5-12 se consideran los promotores más potentes conocidos para permitir una expresión génica robusta en los músculos cardíaco y esquelético. La figura 21 muestra que el vector que comprende el elemento regulador específico de músculo Sk-SH4 tal como se describe en el presente documento operativamente unido al promotor de desmina permite una expresión mucho mejor que los vectores que comprenden el promotor de CMV o el SPc5-12 sin elemento regulador específico de músculo.

De manera similar, cuando el elemento regulador específico de músculo Sk-SH4 se unió operativamente con el promotor de SPc5-12, se observaron mayores niveles de ARNm de luciferasa en el corazón y todos los tipos de músculo esquelético sometidos a prueba (figura 21). Sin embargo, el nivel de ARNm inducido a partir de este constructo fue menor que cuando el elemento regulador específico de músculo Sk-SH4 descrito en el presente documento está operativamente unido al promotor de desmina.

Lista de secuencias

<110> VRIJE UNIVERSITEIT BRUSSEL
VIB vzw
Universiteit Gent

<120> ELEMENTOS REGULADORES DE ÁCIDO NUCLEICO ESPECÍFICOS DE MÚSCULO Y MÉTODOS Y USO DE LOS MISMOS

<130> VUB-062-PCT

<150> 14151960.3

<151> 2014-01-21

<160> 46

<170> PatentIn versión 3.5

<210> 1

<211> 381

<212> ADN

<213> Artificial

<220>

<223> Elemento regulador específico de músculo cardiaco y esquelético 1 (CSk-SH1)

5 <400> 1
 gttctcctct ataaataccc gctctggtat ttgggggttg cagctgttgc tgccagggag 60
 atggttgggt tgacatgcgg ctcttgacaa aacacaaacc cctggtgtgt gtgggcgtgg 120
 gtggtgtgag tagggggatg aatcagggag ggggcggggg acccaggggg caggagccac 180
 acaaagtctg tgcggggggtg ggagcgcaca tagcaattgg aaactgaaag cttatcagac 240
 cctttctgga aatcagccca ctgtttataa acttgaggcc ccaccctcga cagtaccggg 300
 gaggaagagg gcctgcacta gtccagaggg aaactgaggc tcagggctag ctgcccata 360
 gacatacatg gcaggcaggc t 381

<210> 2

<211> 435

10 <212> ADN

<213> Artificial

<220>

<223> Elemento regulador específico músculo cardiaco y esquelético 2 (CSk-SH2)

15 <400> 2
 cgtctataaa ttccagggaa ggtctctgat tggccctgct cattcccagg cccattcctt 60
 gaccagtca ctgaagtcag ggagatgcag taataagact ggctggaatc agggctcttta 120
 ggggtggagg gatggggagg aggcacagca tgtcatcaaa ataaggaaat tgcaaaagaa 180
 agcttgcagg ctactttgaa tgacaatgag aaagacggtg ctgcctgagt gtgttaagga 240
 tccacatggt ctccaaaatc ctccaggagc atacagtcta gtctgggaga tgagacacaa 300
 aaataaccag aacacaacag cttgcactga ctcgagggct ggataagaat atctggaact 360
 ccccatcta ttccagaagc ttgtctcttg gatgaaaatt agacacttaa tgggaaaggg 420
 ctttgaaaag agtgc 435

20 <210> 3

<211> 551

<212> ADN

<213> Artificial

25 <220>

<223> Elemento regulador específico de músculo cardiaco y esquelético 3 (CSk-SH3)

<400> 3

aaaaaataaaa aataaaaaaat aaaaataaat aaagtggatg gcttagagta tcttttctgt	60
ttagacctga ctaaagctta gacataattg ttagtttagg ctctcagggt aaaatttatt	120
actgtaaate caaaaaatcc cttcttcttc tttttttttt ttttttttgt ccttgaatta	180
aatgctgtca cctccttctt gaaaggagaa actatttagtc agatttgaaa atcctcttta	240
tcaccagga aaatcatttt tatggacact ttgtcttctt gtagtctgac ttagaagcag	300
cctgtttttg ataggttgaa gttttcatct tgaacacaaa ccctgtttgt gtgtccccta	360
ctcccagtt tgatgtgcca ggcactttgt tctcaagccc agcagctgtt gtgggatgag	420
gggacatttt gcatgcttag ccagcagctg ccagaaacat ttctaactct gttttggcag	480
gaaatagggc acaagtggaa gccaaagttaa aagaagctgg aaaaataaac agaataactt	540
tagatgtcac t	551

<210> 4
 <211> 430
 <212> ADN
 <213> Artificial

<220>
 <223> Elemento regulador específico de músculo cardíaco y esquelético 4 (CSk-SH4)

<400> 4	
tgactgtag actatttatg tttgcatatt tttcaggatt gtgtacagta aaccttaagt	60
gaatacaaag gatccaattg tcctgtaaga ctcaacttcaa ttacagattg tgctcagtat	120
taaactttgc tagtactttc aggatgcagt caatcaagag gcagggaaag gtctgtcagc	180
atccacagcc tccttttcaa attggacact tagtctctgg tccgaataat aggtgcccc	240
agtgtcacct cacacaatgg ggtatacaca gtttttaaaa atttattttc agctgggcac	300
agtagctcac acctaaaatc ccagcacttt gggaggctga ggcaggtgga tcacctgagg	360
tcaggagttc aagaccagcc tggccaacgt ggcaaaatcc tgtctctact aaaaatacaa	420
aaattaaccg	430

<210> 5
 <211> 454
 <212> ADN
 <213> Artificial

<220>
 <223> Elemento regulador específico de músculo cardíaco y esquelético 5 (CSk-SH5)

<400> 5

cagtttactc accagggatt cagagggcagc actgctgaac cctgagccct tggcacatca 60
 ggttggctgt cagaagtcgg cctttgtaca tacacagttc ctttgtgagg ccagctgcg 120
 tgtcctagga gcggggcctc tctccacagc agagctcagc ctctcaagtg tatggacagc 180
 acgggtgcct gatgggtgga tttagccatg agttgaagggt ggcttgggga gaatgagagt 240
 tctagagata gggagaaggg gttgccaata ggagagtgga attcctgagc acctcgtcac 300
 aggcagccga cagaacatga gccgcagggc ccaggctatt tatacctcgc ctgtcactat 360
 caggggtccc acagctccc ccacctccag ccacacacag caggtccttt tgctctttct 420
 ggtcccttct ctactcctcc ccctccctac ctaa 454

<210> 6
 <211> 453
 <212> ADN
 <213> Artificial

<220>
 <223> Elemento regulador específico de músculo cardíaco y esquelético 6 (CSk-SH6)

<400> 6
 ccctttccct ggcaggatct gccctgtggc ccaaattgggc atgttgccca gggggctccc 60
 tggcactatg ggggaagagt ctctccttcc cctcttatca tctcagttga gtcagacttg 120
 ggggaggggg atacacagtg tgagtcactg ggtacccttt tcctgagctc agcttcatac 180
 cgaggcgatg aggccaaacg ggctggtgac agggacactg agtcaggggc aggggccccg 240
 gtcttactcc tgggcctctg gatttgggac ctacatgagg cttttctatc tgtaaagtca 300
 agcaatggct gggaggcaca cacaaccccc cgcccccccg caggcttctc cttcattggc 360
 ccgggcaagg tcctgtcttc ctctcaggcc gtctctgcac aagcacacac acttcccttc 420
 cctgtccaca ggtggacaat gccctgggct agg 453

<210> 7
 <211> 495
 <212> ADN
 <213> Artificial

<220>
 <223> Elemento regulador específico de músculo esquelético 1 (Sk-SH1)

<400> 7
 ctactcccc gcccaggcag caaggagccc acaccctcat gccctcagc ttcagcccc 60
 acctccagga ggccctacct acgctcatga ccttgctatt ctgggccttg tgtcctgtag 120

ES 2 862 526 T3

	ggagatggac aggagacagc tgggcttcca ggccacccag gcggggggct agccgagggg	180
	agcctgctgg ctctcctgct tgctctaatt tctggggctc cccaaacctt ggcctcagga	240
	gactggggat aggaccggcc ttgaaagtgg ggggaagcttt ggagagccgg gtgctggggt	300
	cttagtgaga tggccagtga aggctgtggt gccccgaggt aagcagggcc tgatcccctc	360
	ctaattcttc agcagcaact ggtgctctga ggctccccct cccccagccc tgccagcctt	420
	cagggacctg ccttccaaag atgggcaggg gagggggacg aggacacca cccactcctc	480
	agaccagcat gtctt	495
5	<210> 8 <211> 344 <212> ADN <213> Artificial	
10	<220> <223> Elemento regulador específico de músculo esquelético 2 (Sk-SH2)	
	<400> 8 ccctccagat gggtttcctg gaatctagat ttcccaggtt ccaaaggaca cccgagtctc	60
	atgcctggaa ctcagtgaga ctaattcacc tctcctctgc cctaattctt atctccagcc	120
	agaagccaac agatcccagg ggactggagc cacaggggct gcacctgttt accgggtatt	180
	tttaggatgg ttgatgaaca cataataccc accctatagt cagagaaaga caatgcctgc	240
	tatgttaatc ctgtggctat tatagtctgt catctcatgg gttggggcag gacactgacc	300
	ctctcagagg ccagagagag gcctcgcaag caggagggtta ggga	344
15	<210> 9 <211> 430 <212> ADN <213> Artificial	
20	<220> <223> Elemento regulador específico de músculo esquelético 3 (Sk-SH3)	
	<400> 9 atggagacaa tccatgaatt cctgagatgc ttggctggta ttagatttta tgggcagctg	60
	cttattctta gggctctgct tctccaaaga cactgaggaa gtccaaagga aacaccagct	120
	ggcgaagagc cacctccagg cccatctgtc catcatcagc ctccaggaat gccagtgtcc	180
	agagggcacc aggtctgcgt ctgtctccct gggatgtgcc ttgtccttg tgggcatttg	240
	gcagtgatca tgcctccctg tctccctcag agatccaaact gtccccattg tggggcccta	300
	ccttccaagg ccggtttaca cctcctgcc aactccgggg cctgccccca gcctgcctca	360
	ctgacaaatg ccagaccaag ggggtccacg tcaggcaaga ggcctcagcc tgtgctctga	420
	caccctcag	430
25	<210> 10 <211> 435 <212> ADN	

<213> Artificial

<220>

<223> Elemento regulador específico de músculo esquelético 4 (Sk-SH4)

5

<400> 10

ttctgagtc	tctaagg	ctcactcca	actcagccc	atgtcctgtc	aattcccact	60
cagtgtctga	tctccttctc	ctcacctttc	ccatctccc	tttgacccaa	gcttcctgag	120
ctctcctccc	attccccctt	ttggagtcct	cctcctctcc	cagaacccag	taataagtgg	180
gctcctccct	ggcctggacc	cccgtggtaa	ccctataagg	cgaggcagct	gctgtctgag	240
gcagggaggg	gctgggtgtg	gaggctaagg	gcagctgcta	agtttaggg	ggctccttct	300
ctcttcttag	agacaacagg	tggctggggc	ctcagtgcc	agaaaagaaa	atgtcttaga	360
ggtatcggca	tgggcctgga	ggagggggga	cagggcaggg	ggaggcacat	tcctcaggac	420
atcgggtcct	agagg					435

<210> 11

10 <211> 474

<212> ADN

<213> Artificial

<220>

15 <223> Elemento regulador específico de músculo esquelético 5 (Sk-SH5)

<400> 11

gactaggaat	aaatcacata	tcctcaatcc	ctggacaact	tgtttacttc	tagtgtagt	60
tttttcttaa	aaaaaaaaatt	gaaatcattc	tgaggctgga	atactttgga	catgcccagc	120
agttcctggc	agttcccaca	gaagcattac	ctcatgactg	gagtgggtaa	agcactactgt	180
gggctatgga	taagactgac	attaaccaca	agcatgtttg	gcagcagact	ggtgctttac	240
aagctccatg	ttcagcagga	gctgcaaagt	gttcctccaa	accaatattt	gtcatttcttg	300
gattctattt	aggaggtcct	gttactcaca	tggtttcaata	tcagcagaag	ccagtttccc	360
tgtgggtaccg	aagtggatcc	tgatgaattt	acccttgtaa	gtaaaaaaaa	tgatgttata	420
cccaaagctt	gaagtacgta	gtggggatgc	cactgaaata	attcagacat	gctt	474

20 <210> 12

<211> 519

<212> ADN

<213> Artificial

25 <220>

<223> Elemento regulador específico de músculo esquelético 6 (Sk-SH6)

<400> 12

gtgctcatag	ctccaccttt	tgttctaat	atggtctttc	cagctccctc	caccccatca	60
------------	------------	-----------	------------	------------	------------	----

	ttgttctcct gggggaacac aggggtgagac gctttgatga actgacatca ccagcaaaaa	120
	aaatatctag caacagctga ggctgatttt agacaatgga aagtggggga gggaagaggt	180
	tctccctgac cctgaaactt tccactcatt ctgggcagct ctatggatgt tttaaaagaa	240
	gaggaagagg ggaggggaaga acattgaaat agagaagtgt actttggcaa ttctaggttg	300
	gcagtttgca tccaggggggt cctggttgcc tttcagcttc ccgtttcaact ctccccaga	360
	ctgtgttgaa tgctgggtcaa actccgttag ttgagtttta gcttttgatt cctgggtattc	420
	aaggagcttg ggcacagggga agaggggagg tcaactcatga tccttaacaa ttctcccaga	480
	tccccagatc aaattgctgt gctattctgg gagtctccg	519
	<210> 13	
	<211> 372	
5	<212> ADN	
	<213> Artificial	
	<220>	
10	<223> Elemento regulador específico de músculo esquelético 7 (Sk-SH7)	
	<400> 13	
	atcgtgtgtc agaggtttgt gtcagcttcc cagcaaggga accagaaagg aaaaggaacc	60
	ggttcctcat gcttcctagg ggaatgcatg catatctgaa gagaaggga tcttatataa	120
	ggctgttttag ctaagggcag ccaccagcca ggtgagcctt acagaagcac agggctgggt	180
	gtctgcagtt ccctagcaga ttaacctggg tcacagtga tcagagctcc agcatgagag	240
	ttccaggtgt ggaactgagc aagtacagat ctgcttttgc tccacttggg agtatatttc	300
	cttcttagtg agcatgggca gcctcctggc caggggaagtc tggcactgtc tgggcctgac	360
	agggaaaccc tg	372
15	<210> 14	
	<211> 20	
	<212> ADN	
	<213> Artificial	
20	<220>	
	<223> Cebador	
	<400> 14	
	cccaccgtcg tattcgtgag	20
25	<210> 15	
	<211> 21	
	<212> ADN	
	<213> Artificial	
30	<220>	
	<223> cebador	
	<400> 15	
35	tcaggcgcat ggtttgtcc c	21
	<210> 16	

<211> 983
<212> ADN
<213> Artificial

5 <220>
<223> Promotor de desmina

<400> 16
accttgcttc ctagctgggc ctttccttct cctctataaa taccagctct ggtatttcgc 60
cttggcagct gttgctgcta gggagacggc tggcttgaca tgcattctct gacaaaacac 120
aaaccctgtg tgtgagtggg tgtgggcggg gtgagtaggg ggatgaatca gagagggggc 180
gagggagaca ggggcgcagg agtcaggcaa aggcgatgcg ggggtgcgac tacacgcagt 240
tggaacagct cgtcagaaga ttctggaaac tatcttgctg gctataaact tgagggaagc 300
agaaggccaa cattcctccc aagggaact gaggtcaga gttaaaacct aggtatcagt 360
gatatgcatg tgccccggcc agggtcactc tctgactaac cggtacctac cctacaggcc 420
tacctagaga ctcttttgaa aggatggtag agacctgtcc gggctttgcc cacagtcgtt 480
ggaaacctca gcattttcta ggcaacttgt gcgaataaaa cacttcgggg gtccttcttg 540
ttcattccaa taacctaaaa cctctcctcg gagaaaatag ggggcctcaa acaaacgaaa 600
ttctctagcc cgctttcccc aggataaggc aggcattcaa atggaaaaaa aggggccggc 660
cgggggtctc ctgtcagctc cttgccctgt gaaaccacgc aggcctgcct gtcttctgtc 720
ctcttggggc tgtccagggg cgcaggcctc ttgcggggga gctggcctcc ccgccccctc 780
gcctgtggcc gcccttttcc tggcaggaca gagggatcct gcagctgtca ggggaggggc 840
gccggggggg gatgtcagga gggctacaaa tagtgcagac agctaagggg ctccgtcacc 900
catcttcaca tccactccag ccggctgccc gcccgctgcc tcctctgtgc gtccgccag 960
ccagcctcgt ccacgccgcc acc 983

10 <210> 17
<211> 1000
<212> ADN
<213> Artificial

15 <220>
<223> Fragmento de elemento regulador de DES

<400> 17
ccaaaagggt caggccaag agggacctgg agtgccaagt ggaggtgtag aggcacggcc 60
agtacccatg gagaatgggt gatgtcctta ggggttagca agtgccgtgt gctaaggagg 120
gggctttgga ggttgggcag gccctctgtg gggctccatt tttgtggggg tgggggctgg 180
agcattatag ggggtgggaa gtgattgggg ctgtcaccct agccttcctt atctgacgcc 240

20

ES 2 862 526 T3

cacccatgcc	tcctcaggtg	ccccctgccc	cccacagctc	ctctcctgtg	ccttgtttcc	300
cagccatgcg	ttctcctcta	taaatacccg	ctctgggtatt	tgggggttggc	agctgttgct	360
gccagggaga	tgggttgggtt	gacatgcggc	tcctgacaaa	acacaaaccc	ctggtgtgtg	420
tgggcgtggg	tgggtgtgagt	agggggatga	atcagggagg	gggcggggga	cccagggggc	480
aggagccaca	caaagtctgt	gcgggggtgg	gagcgcacat	agcaattgga	aactgaaagc	540
ttatcagacc	ctttctggaa	atcagcccac	tgttttataaa	cttgaggccc	caccctcgac	600
agtaccgggg	aggaagaggg	cctgcactag	tccagagggg	aactgaggct	cagggctagc	660
tcgcccatag	acatacatgg	caggcaggct	ttggccagga	tccctccgcc	tgccaggcgt	720
ctccctgccc	tcccttcctg	cctagagacc	cccaccctca	agcctggctg	gtctttgcct	780
gagacccaaa	cctcttcgac	ttcaagagaa	tatttaggaa	caaggtggtt	tagggccttt	840
cctgggaaca	ggccttgacc	ctttaagaaa	tgacccaaaag	tctctccttg	acaaaaaagg	900
ggaccctcaa	actaaagggg	agcctctctt	ctgctgtctc	ccctgacccc	actccccccc	960
accccaggac	gaggagataa	ccagggctga	aagaggcccc			1000

<210> 18
 <211> 1000
 <212> ADN
 <213> Artificial

<220>
 <223> Fragmento de elemento regulador de DES

aatctgggtc	tttgggctgt	aagtgatcaa	agcaatacca	agtcgcttaa	attaaaaaaaa	60
aaaaaaaaaa	aaaaaacgga	aagtttgttg	gcccgaaggcc	aggaaggaca	gtgtgggagc	120
tcaactcaca	aagttgaggg	aagcactgca	ggaaccaagg	ggctggcctg	ctcctcctct	180
ccagcctcct	ctgcttctta	catattgacc	tctctttctt	cctactcccc	cagggggcag	240
gaaacatggc	ttccacaggt	tccagttgaa	gaatcccagt	tccgtctata	aattccaggg	300
aaggtctctg	attggccctg	ctcattccca	ggcccattcc	ttgacccagt	cactgaagtc	360
agggagatgc	agtaataaga	ctggctggaa	tcaggggtctt	taggggtgga	gggatgggga	420
ggaggcacag	catgtcatca	aaataaggaa	attgcaaaaag	aaagcttgca	ggctactttg	480
aatgacaatg	agaaagacgg	tgctgcctga	gtgtgttaag	gatccacatg	gtctccaaaa	540
tcctccagga	gcatacagtc	tagtctggga	gatgagacac	aaaaataacc	agaacacaac	600
agcttgcact	gactcgaggg	ctggataaga	atatctggaa	ctcccccatc	tatttcagaa	660
gcttgtctct	tggatgaaaa	ttagacactt	aatgggaaag	ggctttgaaa	agagtgcagt	720
aacaaagccc	cctttacaat	ttaccgggca	cattcacacc	catcctgagg	ccaaagccac	780
aggctgtgag	gtctcactgt	ctcagcttcc	tgagctataa	aatgggaatg	atgctagtgt	840

ES 2 862 526 T3

	ctacctccta gggttggaga attgggggtc atgggtgtga agtgctcagc agcttggccc	900
	acactaggtg gtcagtacat gtaaggtatt attgttgcta catacattag tagggcctgg	960
	gcctctttaa acctttatag ggtagcatgg caaggctaac	1000
5	<210> 19 <211> 1000 <212> ADN <213> Artificial	
10	<220> <223> Fragmento de elemento regulador de ACTN2	
	<400> 19	
	cagatggatc acctgaagtc aggagtttaa gaccagtctg gccaacatgg caaaaccccc	60
	tctctactaa aaatacaaaa atcagctagg tgtggtggca ggcacctgta ttcctagcta	120
	ctaggggtgg tggggcggga gaatcacttg aactcgggag gcggagattg cggtgagctg	180
	agatcatgcc actgcactcc agcctggaag caagactccg tctcaaaaaa taaaaataaa	240
	aaataaaaaat aaataaagtg gatggccttag agtatctttt ctgtttagac ctgactaaag	300
	cttagacata attgttagtt taggctctca gggtaaaatt tattactgta aatccaaaaa	360
	atcccttctt cttctttttt tttttttttt tggtccttga attaaatgct gtcacctcct	420
	tcttgaaagg agaaactatt agtcagattt gaaaatcctc tttatcaccc aggaaaatca	480
	tttttatgga cactttgtct ttctgtagtc tgacttagaa gcagcctgtt tttgataggt	540
	tgaagttttc atcttgaaca caaacctgtt ttgtgtgtcc cctactcccc agtttgatgt	600
	gccaggcact ttgttctcaa gccagcagc tgttgtggga tgaggggaca ttttgcacgc	660
	ttagccagca gctgccagaa acatttctaa tctggttttg gcaggaaata gggcacaagt	720
	ggaagccaag ttaaaagaag ctggaaaaat aaacagaata acttttagatg tcacttaata	780
	tatgggtccat tttcagccga agatttgccc tagtaatttg ttaatatgac tggactgtga	840
	tcccttccaa aggcagggtt gaatatagtc acctttgaga tccaggatgt agtccagtgc	900
	ttggaatatg cttgtaggag gttattgtta ttattttttc aatgatagac tatactgcaa	960
	atttgtttaa ttgatttaat tagtagctta aaactgttgt	1000
15	<210> 20 <211> 999 <212> ADN <213> Artificial	
20	<220> <223> Fragmento de elemento regulador de ACTN2	
	<400> 20	
	agctgaagcc gtgggacgtc tgagcggcac atttttatgg tcatcacaga tgggtgctggg	60

tttttccttt tttttttttt tttttttttg gtgagaagtc aattggccaa acttgaaatt	120
taccaataaaa accctttaag tgttacaaaa tctactgggca aaggaatata tttgagatgt	180
ggtagcatct gtgaattcca tttacctaag agtatgaatt acaaaccaga tgaatgacag	240
attagttgtc atcaactaca cacagtttat tttttcatgt gagttgcact gtagactatt	300
tatgtttgca tattttttcag gattgtgtac agtaaaccctt aagtgaatac aaaggatcca	360
attgtcctgt aagactcact tcaattacag attgtgctca gtattaaact ttgctagtac	420
tttcaggatg cagtcaatca agaggcaggg aaaggctctgt cagcatccac agcctccttt	480
tcaaattgga cacttagtct ctggtccgaa taataggtgc cccagtgctc acctcacaca	540
atgggggtata cacagttttt aaaaatttat tttcagctgg gcacagtagc tcacacctaa	600
aatcccagca ctttgggagg ctgaggcagg tggatcacct gaggtcagga gttcaagacc	660
agcctggcca acgtggcaaa atcctgtctc tactaaaaat acaaaaatta accgggcatg	720
gtggtggcac ctgtagtccc agctactcgg gaggctgagg caggagaatc ttttgaaccc	780
gggaggcaga ggttgacgtg agccaagatc ccaccactgc actccagcct gggtgacaga	840
gtgagactgt gcctcaaaaa aaattaaaaa aagtattttc atttaaaaag gggcgagtca	900
ccaaccaccc tgaagtctag ttttagtgctc tttttatgtt ggagcttggtg gctgctcctg	960
cattgtcctt ttgaagtccc taagtgtcat cagaattta	999

<210> 21
 <211> 999
 <212> ADN
 <213> Artificial

<220>
 <223> Fragmento de elemento regulador de FLNC

<400> 21	
cacactgacc tcccctggag cgagctcttg agtcagaggg aaagcaaggg aagctacgct	60
gggagttcct tctccctctc ctggcttctc ccggcctctc cacctgaaca gaggcgctca	120
gctggctggg ggacctcaga taatgcctgc aaaacaggac tgtggacat tagggctctgt	180
aaaggggctg tcagctccga gctacctcac ctttaaccctt tctcctcccc agcagctcaa	240
agaggaatgc atcctttctc agatctgtgt ctgagtttac tcaccaggga ttcagaggca	300
gcactgctga accctgagcc cttggcacat caggttggct gtcagaagtc ggcctttgta	360
catacacagt tcccttgtag ggcccagctg cgtgtcctag gagcggggcc tctctccaca	420
gcagagctca gcctctcaag tgtatggaca gcacgggtgc ctgatgggtg gatttagcca	480
tgagttgaag gtggcttggtg gagaatgaga gttctagaga tagggagaag gggttgccaa	540
taggagagtg gaattcctga gcacctcgtc acaggcagcc gacagaacat gagccgcagg	600
gcccaggcta tttatacctc gcctgtcact atcagggtcc ccacagctcc cccacacctc	660

	agccacacac agcaggtcct tttgctcttt ctggtccctt ctctactcct cccctccct	720
	acctaaggac caggcaattt cattcataga ccctgtcat cataagcatt gccccacag	780
	aaccctgaca ctttttcctc agctccgtaa cagggggtgc cctcattggc ccttggggga	840
	aagagcccc aaagatgctc actatccctt cttttacacc aggtctaggg gtagaagagg	900
	aattagatat aatgattcat ctaaaagggc ccagaacctt ccctggctta aagaaacttc	960
	tcaggctcag atatggcctc ctagccagct cagctagac	999
5	<210> 22 <211> 1000 <212> ADN <213> Artificial	
10	<220> <223> Fragmento de elemento regulador de FLNC	
	<400> 22 aaaaaaaaa gaaaaaagaa aaagaaatac gttctctagc tctttctcct caccacaggc	60
	cagagtctgt ttgggtgcct cggtttccct gctgttctat tccccaggcc tgcagccttt	120
	ggcttgcctt gagtgcagct ttgcctgggc atcccagctg gctcccttag gctctccctg	180
	accactctgc ccctggctcct gcctaccac ctgggtcacat ttccccctgg ctgctctggc	240
	tggggctctc cctgtttctg cctggctgac ccaccctttc cctggcagga tctgccctgt	300
	ggcccaaagt ggcattgtgc ccagggggct ccctggcact atgggggaag agtctctcct	360
	tcccccttta tcatctcagt tgagtcagac ttgggggagg gggatacaca gtgtgagtca	420
	ctgggtacct ttttctgag ctacagcttca taccgaggcg atgaggccaa acgggctggg	480
	gacagggaca ctgagtcagg ggcagggggc ccggtcttac tcctgggcct ctggatttgg	540
	gccctacatg aggcttttct atctgtaaag tcaagcaatg gctgggaggc acacacaacc	600
	ccccgcccc ccgcaggctt ctccctcatt ggcccgggca aggtccctgc ttcctctcag	660
	gccgtctctg cacaagcaca cacacttccc ttccctgtcc acaggtggac aatgccctgg	720
	gctaggagcc agccctcgca ggccctgtca acagcctgcc accaactagt ctagaaaatt	780
	ctgtgaacac cagatcagtg ggttttgccc taatcactgg cctcagtttc cccatttacc	840
	tcttcagccc gtgcggggtg tggagggagc atatttctatt cctgacagtt tgggtgtgtg	900
	ggtgtggatg gagaactgtc ggccctcccc gccaccttcc agcgcggcgg caccgggggc	960
	ccaggggggtg ggcgccctca accccgtccc gccgcccggg	1000
15	<210> 23 <211> 1000 <212> ADN <213> Artificial	
20	<220> <223> Fragmento de elemento regulador de ATP2A1	

<400> 23

ccctgtgggc atttgagttt ataacaccac cccattgtg gcacaccct ccacccgta	60
aaacacaggc tctgctcttg gaatcagtct tctgatctg tggctgtgcc ctccaacaga	120
gggcaccctt gggcttccca gctctggggg tagtgggtgc caacaaggag gggcctgggg	180
ctgaagaatc ccacccgctg agctcggcct tctcccttcc ccaactgtcca gctccgcctt	240
tcagcatcct gcctcactcc ccgcccaggc agcaaggagc ccacaccctc atgcccctca	300
gcttcagccc ccacctccag gagggccctac ccacgctcat gaccttgcta ttctgggcct	360
tgtgtcctgt agggagatgg acaggagaca gctgggcttc caggccaccc aggcgggggg	420
ctagccgagg gaagcctgct ggctctcctg cttgctctaa tttctggggc tccccaaacc	480
ttggcctcag gagactgggg ataggaccgg ccttgaaagt gggggaagct ttggagagcc	540
gggtgctggg ttcttagtga gatggccagt gaaggctgtg gtgccccgag gtaagcaggg	600
cctgatcccc tctaattctt ccagcagcaa ctggtgctct gaggctcccc ctccccagc	660
cctgccagcc ttcagggacc tgccctccaa agatgggcag gggaggggga cgaggacacc	720
caccactcc tcagaccagc atgtcttggc tgttggggcc tgagagactt tccctctaag	780
cctttcttta cagatgggta aaccgaagtt ctgcactcat cagggactgg ccagggtgtc	840
tctgtgtccc atgcttttag ctccagccct cagggtgtgac aggaggatca ctttccatcc	900
ctgggcgtgg agaccctgt gggaagggat ccccgagggc gcctctggct cagcctccct	960
ccatggcagt tcacaccac agccttccct agagcagccc	1000

5 <210> 24
<211> 999
<212> ADN
<213> Artificial

10 <220>
<223> Fragmento de elemento regulador de TNNI1

<400> 24

gggccacagt gaaggcccag aggggcaagt agcagctggg ggaagaggac cattgggcta	60
accctgcata tcttataagg caatatggaa atgaggcccc tgaacccatg ccccaaaagc	120
ccaagagggc taaagactca gctttcaggg atttttcagt ggctctggtg cacgtgggat	180
acgaggggtg ctggctgtgg ggagtgatgt gagcagtgtt gcaggctgac ccagtgagt	240
gtttccaggg ctgagcagaa tggcagtggc agcagggcac acctttatgc tcagggaggt	300
ctgaggggtc gcagacactc tgcatcccc tccagatggg tttcctggaa tctagatttc	360
ccaggttcca aaggacaccc gagtctcatg cctggaactc agtgagacta attcacctct	420
cctctgccct aatcttcata tccagccaga agccaacaga tcccagggga ctggagccac	480

aggggctgca cctgtttacc gggatattttt aggatgggtg atgaacacat aatacccacc 540
ctatagtcag agaaagacaa tgctgtctat gttaatcctg tggtattat agtctgtcat 600
ctcatgggtt ggggcaggac actgacctc tcagaggcca gagagaggcc tcgcaagcag 660
gaggttaggg agccccagcc atgctcccca tttggagaag gagagagtaa tgggcagggg 720
tgtcctaagg agacacaccc cagtgtccac agaggtaaag gccctgccct ctctacaacg 780
acattctgct gcctaataga ggttgctcta attctgttct ggtcgggtca actttgagaa 840
gcctctatgt gagacggctg ccaggagaaa atgacagggt ctggggaggc ataggggttg 900
gggacatagg attccaagtg gagaggctgt gatggataat gaaggcaggc tccagccaaa 960
aggagcagct agaggggaga aggaatttag gggctttgc 999

<210> 25
<211> 999
5 <212> ADN
<213> Artificial

<220>
10 <223> Fragmento de elemento regulador de TNNI1

<400> 25
ttgtggggtc tgggaagtgg ggaattaaaa gacacttgta aagtgcctgc caataattgt 60
atcccagagc cacccttcc tctgccctaa ttctctcctt tccctttgtg cccatgaagc 120
actatcccaa cccaatgaa cacagctttg gccacctccc ctgaatctcc caacgggtca 180
gaaagccttc tgggcctccc tccacctctg cctccgccac caggggaggg taaggctggg 240
gtggaggagg tggaggtgta gtgtgacccc agagtctcag ctgtcatgct ggggcacaga 300
gggaggggta tggagacaat ccatgaattc ctgagatgct tggctggtat tagattttat 360
gggcagctgc ttattcttag ggctctgctt ctccaaagac actgaggaag tccaaaggaa 420
acaccagctg gcgaagagcc acctccaggc ccatctgtcc atcatcagcc tccaggaatg 480
ccagtgtcca gagggcacca ggtctgcgtc tgtctccctg ggatgtgcct tgtccttggg 540
gggcatttgg cagtgatcat gcctccctgt ctccctcaga gatccaactg tccccattgt 600
ggggccctac cttccaaggc cggtttacac ctccctgcaa gctccggggc ctgccccag 660
cctgcctcac tgacaaatgc cagaccaagg ggtccacgt caggcaagag gcctcagcct 720
gtgctctgac acccctcaga cgggggccct tgccaggctc tgtggaagac aagcgaggac 780
tgataagtca ggatgaagat agcagccact ggaaggcttg gagagccagg attccatctc 840
ctctacaagt agcaagcaag gagataatga gaaaatcctc ccccgaggg ggggagcaaa 900
gagtctctgg tggttgcctc agagctgggc tttcccttcc ttgaggcttt gcagcgtgga 960
agggactatc cgctagacta tgagattgag tctgtgtgt 999

<210> 26

<211> 1000
<212> ADN
<213> Artificial

5 <220>
<223> Fragmento de elemento regulador de MYLPF1

<400> 26
tctttttgggg atccaaggcc ccagtcccgt caacccctct cccctcactg ccacatccct 60
ctgtgctcct catgttccac tgagaatccc gacctcaatc tcatgagcgc tccagggtgt 120
ccttagatca caagcataaa gacctgacgt tatgagctcc aaccttccaa cttctaaacg 180
ccctcttctt tctctgaccc ccatattctg atttccatat agcatccacc atcctggaat 240
tccagtgaga cctagctccc gacttctcag ctccctccttg gtttctgagt cctctaaggt 300
ccctcactcc caactcagcc ccagtgcctg tcaattccca ctcaagtgtc gatctccttc 360
tcctcacctt tcccatctcc cgtttgaccc aagcttcctg agctctcctc ccattcccct 420
ttttggagtc ctccctctct cccagaaccc agtaataagt gggctcctcc ctggcctgga 480
cccccggtgt aaccctataa ggcgaggcag ctgctgtctg aggcaggag gggctggtgt 540
gggaggctaa gggcagctgc taagtttagg gtggctcctt ctctcttctt agagacaaca 600
ggtggctggg gcctcagtgc ccagaaaaga aaatgtctta gaggtatcgg catgggcctg 660
gaggaggggg gacagggcag ggggaggcat ctccctcagg acatcgggtc ctagaggag 720
cgggaggaga aggagatggt tgtccttgcc aacttggggc ttccctcagcc actatttttc 780
cagacttctg ctgcatggag gggactgggt cactgaggcc cagagggaga agagagggt 840
catcaaagtc acacaatcag ccaagctggc tcttgccag aataaagtga gctgccactg 900
gctatcaagg cacctcacag aaagtgaaca gctggctgtc cttttagggt cttttcctcc 960
cctcttgagt tccagcctcc acaccagtac ccagagaaaa 1000

<210> 27
<211> 999
<212> ADN
<213> Artificial

15 <220>
<223> Fragmento de elemento regulador de MYH1

<400> 27
gggatcatta gctaaatctc actggcctct tgagatttca actgaatgtt cctgtgaagt 60
accacatatt atgtctctac aaacactatt gcagttcctc ctgcctcctg ggaaaaaagt 120
gaaatatgct tggctcaaga tattcacttc tcatcacaga gatccttgac atttgggaata 180
taattgagtc ctgagactcc ctctgcaggg tcctgtcgcc atataataga gggactacta 240
aaagcaacag tattaaagct cagactagga ataaatcaca taccctcaat ccctggacaa 300

cttgtttact tctagtgtta gttttttctt aaaaaaaaaa ttgaaatcat tctgaggctg 360
gaatactttg gacatgcccc gcagttcctg gcagttccca cagaagcatt acctcatgac 420
tggagtgggt aaagcatact gtgggctatg gataagactg acattaacca caagcatgtt 480
tggcagcaga ctggtgcttt acaagctcca tgttcagcag gagctgcaaa gtgttcctcc 540
aaaccaatat ttgtcattct tggattctat ttaggaggtc ctgttactca catgtttcaa 600
tatcagcaga agccagtttc cctgtggtac cgaagtggat cctgatgaat ttacccttgt 660
aagtaaaaaa aatgatgtta tacccaaagc ttgaagtacg tagtggggat gccactgaaa 720
taattcagac atgctttctc tggcatctaa aggagagaaa gaccctttgg caaccaagag 780
acttacaag cgagaggagt tgtcattcct caggtcttg gcgttgccaa aggctccag 840
tagggggttg gcaactgatga tttgatcttc cagagtcccc tgcaaaggca agagcagtc 900
ttgcatctgg ggcttgggaa tttcctacct gagagtcccg acagagcctg gattctgact 960
aacaatcaga ctcacctgca ttttgccaga agtaacttc 999

<210> 28
<211> 1000
5 <212> ADN
<213> Artificial

<220>
10 <223> Fragmento de elemento regulador de TPM3

<400> 28
gtgcctacct gtttacttct ttcttctgcc tgcttctgct cagcttcagc ttgctctgcc 60
cgatccagag cattctcctt gtctaacttc agcatctgca tctttttctt gatggcctcc 120
atcatgagca gtggctgttg gtaggctcac ctgtgaacac tggagaactg gagactgggg 180
caagaaagaa ggggctgctg cctgagtgac caggaggtcc ccagacttga gtctttatct 240
gtgctcatag ctccaccttt tgttccta atgggtcttc cagctccctc caccocatca 300
ttgttctcct gggggaacac aggggtgagac gctttgatga actgacatca ccagcaaaaa 360
aaatatctag caacagctga ggctgatttt agacaatgga aagtggggga gggaagaggt 420
tctccctgac cctgaaactt tccactcatt ctgggcagct ctatggatgt tttaaaagaa 480
gaggaagagg ggaggggaaga acattgaaat agagaagtgt actttggcaa ttctaggttg 540
gcagtttgca tccaggggggt cctggttgcc tttcagcttc ccgtttcact ctccccaga 600
ctgtgttgaa tgctgggtcaa actccgttag ttgagtttta gcttttgatt cctggtattc 660
aaggagcttg ggcacaggga agaggggagg tcaactcatga tccttaacaa ttctcccaga 720
tccccagatc aaattgctgt gctattctgg gactctccga ttggcaggga gacgtttttc 780
ctccccacca agagccacaa gagtagagta agaggtaggg ttggaatctc caaccctatc 840

	cttgaaggct atcccagagc ttcaagtggg gtggggagag aaacagggga gggtcagaac	900
	aatccagaca ggatcatacc ctttgttttc ccacagtaat cttaaaatag aactgttgta	960
	cccacacaga ctttgggtctg tgtggctctc tcctcttctt	1000
	<210> 29	
	<211> 999	
5	<212> ADN	
	<213> Artificial	
	<220>	
	<223> Fragmento de elemento regulador de ANKRD2	
10	<400> 29	
	ggcatgcaat ctccagcaga gtgagcctgg aactgccagt cttctgggcc ctgatcccag	60
	ctcccctggg gtccctcatga gtccctcagga aaagccactc tttgcttttc tttaaagcag	120
	agccatgctg cctgccctgc ctccctccatc agttgctggg aggattaagc aggaagacag	180
	atgtgaaagt gcttgcaagg cgaagacagc acctcaggaa ggtgggtctgc tcctgcagca	240
	agaggggcta acttccggcc ctggagacgc tgtagccatc gtgatgttaa agcaaaatct	300
	ctgacagatt tagatcgtgt gtcagagggt tgtgtcagct tcccagcaag ggaaccagaa	360
	aggaaaagga accggttcct catgcttcct aggggaatgc atgcatactt gaagagaagg	420
	gaatcttata taaggctgtt tagctaaggg cagccaccag ccaggtgagc cttacagaag	480
	cacagggctg ggtgtctgca gttccctagc agattaacct gggtcacagt gactcagagc	540
	tccagcatgc gagttccagg tgtggaactg agcaagtaca gatctgcttt tgctccactt	600
	gggagtattt ttccttctta gtgagcatgg gcagcctcct ggccagggaa gtctggcact	660
	gtctgggcct gacagggaaa ccctgggagg gtagaaggat ccagagtagc tgctgttcct	720
	cgctagctgg gcttagtgct ttccggagac ccctttcttg aagcaagact ctgtaagccc	780
	tgcaagagtc ccctgagcac ttaccaagag aggagacaga ttggataagg tttgttcatg	840
	actaaaagtc acccagccaa atgcagtgat gcttaacggg gaagcatgtg ggaccccgga	900
	gagccgaagg ccagcgtgag ctgtgatcag agagggagac agctgccctt cctgctacca	960
	cggccctggc ctggacaagt agagtgtgac cctcctcac	999
	<210> 30	
15	<211> 663	
	<212> ADN	
	<213> Artificial	
	<220>	
20	<223> Promotor de CMV	
	<400> 30	
	acgcgtggag ctagttatta atagtaatca attacggggc cattagttca tagcccatat	60
	atggagttcc gcgttacata acttacggta aatggcccgcc ctggctgacc gccaacgac	120

ccccgcccacat tgacgtcaat aatgacgtat gttcccatag taacgtcaat agggactttc 180
cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg 240
tatcatatgc caagtacgcc ccctattgac gtcaatgacg gtaaagtggc cgcctggcat 300
tatgcccagt acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc 360
atcgctatta ccatggtgat gcgggttttg cagtacatca atgggcgtgg atagcggttt 420
gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt gttttgcacc 480
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 540
gtaggcgtgt acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg 600
cctgggagacg ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc 660
tcc 663

5 <210> 31
<211> 20
<212> ADN
<213> Artificial

10 <220>
<223> Cebador

<400> 31
tgtgtccgtc gtggatctga 20

15 <210> 32
<211> 22
<212> ADN
<213> Artificial

20 <220>
<223> cebador

<400> 32
gcctgcttca ccaccttctt ga 22

25 <210> 33
<211> 20
<212> ADN
<213> Secuencia artificial

30 <220>
<223> cebador

<400> 33
gtccctcact cccaactcag 20

35 <210> 34
<211> 23
<212> ADN
<213> Secuencia artificial

40 <220>
<223> cebador

<400> 34
gaggagaagg agatcagaca ctg 23

45

5 <210> 35
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> cebador

10 <400> 35
 tagctgggcc ttctctctc 20

15 <210> 36
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> cebador

20 <400> 36
 cgtctcccta gcagcaacag 20

25 <210> 37
 <211> 171
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Fragmento a de elemento regulador específico de músculo esquelético 4 (Sk-SH4a)

30 <400> 37
 ctctaaggtc cctcactccc aactcagccc catgtcctgt caattccac tcagtgtctg 60
 atctccttct cctcaccttt cccatctccc gtttgaccca agcttcctga gctctcctcc 120
 cattccccctt ttggaggtcc tcctcctctc ccagaaccca gtaataagtg g 171

35 <210> 38
 <211> 51
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <223> Fragmento b de elemento regulador específico de músculo esquelético 4 (Sk-SH4b)

45 <400> 38
 tggcctggac ccccggtgta accctataag gcgaggcagc tgctgtctga g 51

50 <210> 39
 <211> 60
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Fragmento c de elemento regulador específico de músculo esquelético 4 (Sk-SH4c)

55 <400> 39
 gcagggaggg gctggtgtgg gaggctaagg gcagctgcta agtttaggt ggctcctct 60

60 <210> 40
 <211> 41
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Fragmento d de elemento regulador específico de músculo esquelético 4 (Sk-SH4d)

5 <400> 40
 aggagggggg acagggcagg gggaggcatc ttctcagga c 41

<210> 41
 <211> 120
 10 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Fragmento e de elemento regulador específico de músculo esquelético 4 (Sk-SH4e)

15 <400> 41
 gcagggaggg gctgggtgtgg gaggctaagg gcagctgcta agtttagggt ggctccttct 60
 ctcttcttag agacaacagg tggctggggc ctcagtgcc agaaaagaaa atgtcttaga 120

<210> 42
 20 <211> 17
 <212> ADN
 <213> Secuencia artificial

<220>
 25 <223> cebador

<400> 42
 gtgcctact acatcaa 17

30 <210> 43
 <211> 16
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> cebador

<400> 43
 aggttgct ggtcca 16

40 <210> 44
 <211> 8069
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Constructo de plásmido AAVss-Sk-SH4-Des-MVM-MD1

<400> 44

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc	60
gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca	120
actccatcac taggggttcc tcgtacgttc tgagtccctc aaggccctc actcccaact	180
cagcccatg tcctgtcaat tcccactcag tgtctgatct ccttctcctc acctttccca	240
tctcccgttt gacccaagct tcctgagctc tcctccatt cccctttttg gagtccctcct	300
cctctcccag aaccagtaa taagtgggct cctccctggc ctggaccccc gtggtaaccc	360
tataaggcga ggcagctgct gtctgaggca gggaggggct ggtgtgggag gctaagggca	420
gctgctaagt ttaggggtggc tccttctctc ttcttagaga caacaggtgg ctggggcctc	480
agtgcccaga aaagaaaatg tcttagaggt atcggcatgg gcctggagga ggggggacag	540
ggcaggggga ggcattcttc tcaggacatc gggtcctaga gggaccttgc ttcctagctg	600
ggcctttcct tctcctctat aaataccagc tctggtatct cgccttggca gctgttgctg	660
ctagggagac ggctggcttg acatgcatct cctgacaaaa cacaaacccg tgggtgtgagt	720
gggtgtgggc ggtgtgagta gggggatgaa tcagagaggg ggcgagggag acaggggcgc	780
aggagtccag caaaggcgat gcgggggtgc gactacacgc agttggaaac agtcgtcaga	840
agattctgga aactatcttg ctggctataa acttgaggga agcagaaggc caacattcct	900
cccaaggga actgaggctc agagttaaaa cccagggtatc agtgatatgc atgtgccccg	960
gccagggcca ctctctgact aaccggtaac taccctacag gcctacctag agactctttt	1020
gaaaggatgg tagagacctg tccgggcttt gccacagtc gttggaaacc tcagcatttt	1080
ctaggcaact tgtgcgaata aaacacttgc ggggtccttc ttgttcattc caataaccta	1140
aaacctctcc tcggagaaaa tagggggcct caaacaaacg aaattctcta gcccgctttc	1200
cccaggataa ggcaggcatc caaatggaaa aaaaggggccc ggccgggggt ctcctgtcag	1260
ctccttgccc tgtgaaaccc agcaggcctg cctgtcttct gtcctcttgg ggctgtccag	1320
gggcgcaggc ctcttgccgg ggagctggcc tccccgcccc ctgcctgtg gccgcccttt	1380
tcctggcagg acagagggat cctgcagctg tcaggggagg ggcgccgggg ggtgatgtca	1440
ggagggctac aaatagtga gacagctaag gggctccgtc acccatcttc acatccactc	1500
cagccggctg cccgccgct gcctcctctg tgcgtccgcc cagccagcct cgtccacgcc	1560
gccacctcta gaaagaggta agggtttaag ggatggttgg ttggtggggg attaattgtt	1620

aattacctgg	agcacctgcc	tgaatcact	ttttttcagg	ttggacgcgt	gccaccatgc	1680
tgtggtggga	ggaagtggag	gactgctacg	agagagagga	cgtgcagaag	aaaaccttca	1740
ccaagtgggt	gaacgcccag	ttcagcaagt	tcggcaagca	gcacatcgag	aacctgttca	1800
gcgacctgca	ggatggcagg	agactgctgg	atctgctgga	gggactgacc	ggccagaagc	1860
tgcccaggga	gaagggcagc	accagagtgc	acgccctgaa	caacgtgaac	aaggccctga	1920
gagtgctgca	gaacaacaac	gtggacctgg	tgaatatcgg	cagcaccgac	atcgtggacg	1980
gcaaccacaa	gctgaccctg	ggcctgatct	ggaacatcat	cctgcactgg	caggtgaaga	2040
acgtgatgaa	gaacatcatg	gccggcctgc	agcagaccaa	cagcgagaag	atcctgctga	2100
gctgggtgag	gcagagcacc	agaaactacc	cccaggtgaa	cgtgatcaac	ttcaccacct	2160
cctggagcga	cggcctggcc	ctgaacgccc	tgatccacag	ccacagaccc	gacctgttcg	2220
actggaacag	cgtggtgtgt	cagcagagcg	ccaccagag	actggagcac	gccttcaaca	2280
tcgccagata	ccagctgggc	atcgagaagc	tgctggaccc	cgaggacgtg	gacaccacct	2340
accccgacaa	gaaaagcatc	ctgatgtata	ttacctctct	gtttcaggtg	ctgccccagc	2400
aggtgtccat	cgaggccatc	caggaagtgg	aaatgctgcc	caggcccccc	accgtgtccc	2460
tggcccaggg	ctatgagaga	accagcagcc	ccaagcccag	attcaagagc	accgtgtccc	2520
tggcccaggg	ctatgagaga	accagcagcc	ccaagcccag	attcaagagc	tacgcctaca	2580
cccaggccgc	ctacgtgacc	acctccgacc	ccaccagaag	ccccttcccc	agccagcacc	2640
tggaggcccc	cgaggacaag	agcttcggca	gcagcctgat	ggagagcgaa	gtgaacctgg	2700
acagatacca	gaccgcctctg	gaggaagtgc	tgtcttggct	gctgtccgcc	gaggacaccc	2760
tgcaggccca	gggcgagatc	agcaacgacg	tggaagtggg	gaaggaccag	ttccacaccc	2820
acgagggcta	catgatggat	ctgaccgccc	accagggcag	agtgggcaat	atcctgcagc	2880
tgggcagcaa	gctgatcggc	accggcaagc	tgagcgagga	cgaggagacc	gaagtgcagg	2940
agcagatgaa	cctgctgaac	agcagatggg	agtgcctgag	agtggccagc	atggagaagc	3000
agagcaacct	gcaccgcgtg	ctgatggacc	tgcagaacca	gaagctgaag	gagctgaacg	3060
actggctgac	caagaccgag	gagcggacca	gaaagatgga	ggaggagccc	ctgggccccg	3120
acctggagga	cctgaagaga	caggtgcagc	agcacaaagt	gctgcaggag	gacctggaac	3180
aggagcaggt	gcgcgtgaac	agcctgaccc	acatggtggg	cgtggtggac	gagagcagcg	3240
gcgaccacgc	cacagccgcc	ctggaagagc	agctgaaagt	gctgggcgac	agatgggcca	3300
acatctgccg	gtggaccgag	gacagatggg	tgctgctgca	ggacatcctg	ctgaagtggc	3360
agagactgac	agaggagcag	tgctgttcta	gcgcctggct	gagcgagaag	gaggacgccg	3420
tgaacaagat	ccacaccacc	ggcttcaagg	accagaacga	gatgctgagc	agcctgcaga	3480
agctggccgt	gctgaaggcc	gatctggaga	agaaaaagca	gagcatgggc	aagctgtact	3540

ccctgaagca	ggacctgctg	tccaccctga	agaacaagag	cgtgaccag	aaaaccgagg	3600
cctggctgga	caatttcgcc	cggtgctggg	acaatctggt	gcagaaactg	gagaagagca	3660
ccgcccagat	cagccaggcc	gtgaccacca	cccagcccag	cctgacacag	accaccgtga	3720
tggagaccgt	gaccacagt	accaccagg	agcagatcct	ggtgaagcac	gcccaggagg	3780
agctgcccc	tccccccct	cagaagaagc	ggcagatcac	agtggacacc	ctggagagac	3840
tgcaggagct	gcagggaagcc	accgacgagc	tggacctgaa	gctgagacag	gccgaagtga	3900
tcaagggcag	ctggcagcct	gtgggcgcatc	tgtgatcga	cagcctgcag	gaccacctgg	3960
agaaagtga	ggccctgcgg	ggcgagatcg	ccccctgaa	ggagaatgtg	agccacgtga	4020
acgacctggc	cagacagctg	accacctgg	gcatccagct	gagccctac	aatctgagca	4080
ccctggaaga	tctgaacacc	cggtggaaac	tgtgcaggt	ggccgtggag	gatagagtga	4140
ggcagctgca	cgaggcccac	agagacttcg	gccctgcctc	ccagcacttc	ctgagcacca	4200
gcgtgcagg	cccctgggag	agagccatct	cccccaaca	agtgccctac	tacatcaacc	4260
acgagaccca	gaccacctgc	tgggaccacc	ctaagatgac	cgagctgtac	cagagcctgg	4320
ccgacctgaa	caatgtgcgg	ttcagcgct	acagaaccgc	catgaagctg	cggagactgc	4380
agaaggccct	gtgcctggac	ctgctgagcc	tgagcgccgc	ctgcgacgcc	ctggaccagc	4440
acaacctgaa	gcagaacgac	cagcccatgg	acattctgca	gatcatcaac	tgccctgacca	4500
ccatctacga	tcggtggag	caggagcaca	acaacctggt	gaacgtgccc	ctgtgcgtgg	4560
acatgtgcct	gaattggctg	ctgaacgtgt	acgacaccgg	caggaccggc	agaatcagag	4620
tgtgtcctt	caagaccggc	atcatcagcc	tgtgcaaggc	ccacctggag	gataagtacc	4680
gctacctgtt	caagcaggtg	gccagcagca	ccggcttctg	cgatcagagg	agactgggcc	4740
tgctgtgca	cgatagcatc	cagatcccta	ggcagctggg	cgaagtggcc	agctttggcg	4800
gcagcaacat	cgagccctct	gtgaggagct	gcttccagtt	cgccaacaac	aagcccgaga	4860
tgcaggccgc	cctgttcctg	gattggatga	ggctggagcc	ccagagcatg	gtgtggctgc	4920
ctgtgtgca	cagagtggcc	gccgccgaga	ccgccaaagca	ccaggccaag	tgcaacatct	4980
gcaaggagt	ccccatcatc	ggcttccgg	acaggagcct	gaagcacttc	aactacgaca	5040
tctgccagag	ctgctttttc	agcggcagag	tggccaagg	ccacaagatg	cactacccca	5100
tggtggagta	ctgcaccccc	accacctccg	gcgaggatgt	gagagacttc	gccaaagtgc	5160
tgaagaataa	gttccggacc	aagcggctact	ttgccaagca	ccccaggatg	ggctacctgc	5220
ccgtgcagac	cgtgctggag	ggcgacaaca	tggagaccga	caccatgtga	tgatgactcg	5280
agaataaaag	atctttatct	tcattagatc	tgtgtgttgg	ttttttgtgt	gaggaacccc	5340
tagtgatgga	gttggccact	ccctctctgc	gcgctcgctc	gctcactgag	gccgggcgac	5400

caaaggtcgc	cgcagccccg	ggctttgccc	gggcggcctc	agtgagcgag	cgagcgcgca	5460
gctgcctgca	ggggcgcctg	atgcggtatt	ttctccttac	gcctctgtgc	ggtatttcac	5520
accgcatacg	tcaaagcaac	catagtacgc	gccctgtagc	ggcgcatata	gcgcggcggg	5580
tgtggtggtt	acgcgcagcg	tgaccgctac	acttgccagc	gccctagcgc	ccgtcccttt	5640
cgctttcttc	ccttcctttc	tcgccacggt	cgccggcttt	ccccgtcaag	ctctaaatcg	5700
ggggctccct	ttagggttcc	gatttagtgc	tttacggcac	ctcgacccca	aaaaacttga	5760
tttggtgat	ggttcacgta	gtgggccatc	gccctgatag	acggtttttc	gccctttgac	5820
gttgagttcc	acgttcttta	atagtggact	cttggtccaa	actggaacaa	cactcaaccc	5880
tatctcgggc	tattcttttg	atttataagg	gattttgccg	atttcggcct	attggttaaa	5940
aaatgagctg	atttaacaaa	aatttaacgc	gaattttaac	aaaatattaa	cgtttacaat	6000
tttatggtgc	actctcagta	caatctgctc	tgatgccgca	tagttaagcc	agccccgaca	6060
cccgccaaca	cccgctgacg	cgccctgacg	ggcttgtctg	ctcccggcat	ccgcttacag	6120
acaagctgtg	accgtctccg	ggagctgcat	gtgtcagagg	ttttcacctg	catcacccga	6180
acgcgcgaga	cgaaggggcc	tcgtgatacg	cctattttta	taggttaatg	tcattgataat	6240
aatggtttct	tagacgtcag	gtggcacttt	tcggggaaat	gtgcgcggaa	cccctatttg	6300
tttatttttc	taaatacatt	caaatatgta	tccgctcatg	agacaataac	cctgataaat	6360
gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	catttcctgt	tcgcccttat	6420
tccctttttt	goggcatttt	gccttcctgt	ttttgctcac	ccagaaacgc	tggtgaaagt	6480
aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	atcgaaactg	atctcaacag	6540
cggttaagatc	cttgagagtt	ttcgccccga	agaacgtttt	ccaatgatga	gcacttttta	6600
agttctgcta	tgtggcgcg	tattatcccc	tattgacgcc	gggcaagagc	aactcggtcg	6660
ccgcatacac	tattctcaga	atgacttgg	tgagtactca	ccagtcacag	aaaagcatct	6720
tacggatggc	atgacagtaa	gagaattatg	cagtgtctgc	ataaccatga	gtgataaac	6780
tgccggccaac	ttacttctga	caacgatcgg	aggaccgaag	gagctaaccg	cttttttgca	6840
caacatgggg	gatcatgtaa	ctcgccctga	tcgttgggaa	ccggagctga	atgaagccat	6900
accaaacgac	gagcgtgaca	ccacgatgcc	tgtagcaatg	gcaacaacgt	tcgcgaaact	6960
attaactggc	gaactactta	ctctagcttc	ccggcaacaa	ttaatagact	ggatggaggc	7020
ggataaaagt	gcaggaccac	ttctgcgctc	ggcccttccg	gctggctggt	ttattgctga	7080
taaatctgga	gccgggtgag	gtgggtctcg	cggtatcatt	gcagcactgg	ggccagatgg	7140
taagccctcc	cgtatcgtag	ttatctacac	gacggggagt	caggcaacta	tggtgaacg	7200
aaatagacag	atcgctgaga	taggtgcctc	actgattaag	cattggtaac	tgctcagacca	7260
agtttactca	tatatacttt	agattgattt	aaaacttcat	ttttaattta	aaaggatcta	7320

ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca 7380
 ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg 7440
 cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga 7500
 tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa 7560
 tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc 7620
 tacatacctc gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg 7680
 tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggc cgggctgaac 7740
 ggggggttcg tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct 7800
 acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc 7860
 ggtaagcggc agggctcgaa caggagagcg cacgaggag cttccagggg gaaacgcctg 7920
 gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg 7980
 ctctcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct 8040
 ggcccttttg tggccttttg ctccatgt 8069

<210> 45
 <211> 7217
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Constructo de plásmido AAVss-SkSH4-Des-MVM-FST-2A-Luc2

<400> 45
 cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc tcgtacgttc tgagtccctc aaggtccctc actcccaact 180
 cagcccatg tcctgtcaat tcccactcag tgtctgatct cttctcctc acctttccca 240
 tctcccgttt gacccaagct tcctgagctc tcctcccatt cccctttttg gagtccctcct 300
 cctctcccag aaccagtaa taagtgggct cctccctggc ctggaccccc gtggtaaccc 360
 tataaggcga ggcagctgct gtctgaggca gggaggggct ggtgtgggag gctaagggca 420
 gctgctaagt ttaggggtggc tccttctctc ttcttagaga caacagggtg ctggggcctc 480
 agtgcccaga aaagaaaatg tcttagaggt atcggcattg gcctggagga ggggggacag 540
 ggcaggggga ggcattcttc tcaggacatc gggtcctaga gggacctgc ttcttagctg 600
 ggcctttcct tctcctctat aaataccagc tctgggtattt cgccttggca gctgttgctg 660
 ctagggagac ggctggcttg acatgcatct cctgacaaaa cacaaaccg tgggtgtgagt 720
 ggggtgtgggc ggtgtgagta gggggatgaa tcagagaggg ggcgaggag acaggggcgc 780

aggagtcagg	caaaggcgat	gcgggggtgc	gactacacgc	agttggaaac	agtcgtcaga	840
agattctgga	aactatcttg	ctggctataa	acttgagggg	agcagaaggc	caacattcct	900
cccaagggaa	actgaggctc	agagttaaaa	cccaggtatc	agtgatatgc	atgtgccccg	960
gccaggggtca	ctctctgact	aaccggtacc	taccctacag	gcctacctag	agactctttt	1020
gaaaggatgg	tagagacctg	tccgggcttt	gccacacgtc	gttggaacc	tcagcatttt	1080
ctaggcaact	tgtgcgaata	aaacacttcg	ggggtccttc	ttgttcattc	caataaccta	1140
aaacctctcc	tgggagaaaa	tagggggcct	caaacaaacg	aaattctcta	gcccgttttc	1200
cccaggataa	ggcaggcatc	caaatggaaa	aaaagggggc	ggccgggggt	ctcctgtcag	1260
ctccttgccc	tgtgaaaccc	agcaggcctg	cctgtcttct	gtcctcttgg	ggctgtccag	1320
ggcgcgaggc	ctcttgcggg	ggagctggcc	tccccgcccc	ctcgcctgtg	gccgcccttt	1380
tcctggcagg	acagagggat	cctgcagctg	tcaggggagg	ggcgccgggg	ggtgatgtca	1440
ggagggctac	aaatagtgca	gacagctaag	gggctccgtc	acccatcttc	acatccactc	1500
cagccggctg	cccgcccgct	gcctcctctg	tgcgtccgcc	cagccagcct	cgtccacgcc	1560
gccacctcta	gaaagaggta	agggtttaag	ggatggttgg	ttggtgggg	attaatgttt	1620
aattacctgg	agcacctgcc	tgaatcact	ttttttcagg	ttggacgcgt	atggtccgcg	1680
cgaggcacca	gccgggtggg	ctttgcctcc	tgctgctgct	gctctgccag	ttcatggagg	1740
accgcagtgc	ccaggctggg	aactgctggc	tccgtcaagc	gaagaacggc	cgctgccagg	1800
tcctgtacaa	gaccgaactg	agcaaggagg	agtgtctcag	caccggccgg	ctgagcacct	1860
cgtggaccga	ggaggacgtg	aatgacaaca	cactcttcaa	gtggatgatt	ttcaacgggg	1920
gcgcccccaa	ctgcatcccc	tgtaaagaaa	cgtgtgagaa	cgtggactgt	ggacctggga	1980
aaaaatgccg	aatgaacaag	agaacaaac	cccgtgcgt	ctgcgccccg	gattgttcca	2040
acatcacctg	gaaggggtcca	gtctgcgggc	tggatgggaa	aacctaccgc	aatgaatgtg	2100
cactcctaaa	ggcaagatgt	aaagagcagc	cagaactgga	agtccagtac	caaggcagat	2160
gtaaaaagac	ttgtcgggat	gttttctgtc	caggcagctc	cacatgtgtg	gtggaccaga	2220
ccaataatgc	ctactgtgtg	acctgtaatc	ggatttgccc	agagcctgct	tcctctgagc	2280
aatatctctg	tgggaatgat	ggagtcacct	actccagtgc	ctgccacctg	agaaaggcta	2340
cctgcctgct	gggcagatct	attggattag	cctatgaggg	aaagtgtatc	aaagcaaagt	2400
cctgtgaaga	tatccagtgc	actggtggga	aaaaatgttt	atgggatttc	aaggttggga	2460
gaggccgggtg	ttccctctgt	gatgagctgt	gccctgacag	taagtcggat	gagcctgtct	2520
gtgccagtga	caatgccact	tatgccagcg	agtgtgccat	gaaggaagct	gcctgctcct	2580
caggtgtgct	actggaagta	aagcactccg	gatcttgcaa	ctccatttcg	gaagacaccg	2640
aggaagagga	ggaagatgaa	gaccaggact	acagctttcc	tatatcttct	attctagagt	2700

gggtcgagag	gtccggcggc	ggagagggca	gaggaagtct	tctaacatgc	ggtgacgtgg	2760
aggagaatcc	cggcccaatg	gaagatgcc	aaaacattaa	gaagggccca	gcgccattct	2820
accactcga	agacgggacc	gccggcgagc	agctgcacaa	agccatgaag	cgctacgccc	2880
tggtgcccgg	caccatcgcc	tttaccgacg	cacatatcga	ggtggacatt	acctacgccg	2940
agtacttcga	gatgagcgtt	cggctggcag	aagctatgaa	gcgctatggg	ctgaatacaa	3000
accatcggat	cgtgggtgtg	agcgagaata	gcttgcagtt	cttcatgccc	gtgttgggtg	3060
ccctgttcat	cgtgtgggt	gtggccccag	ctaacgacat	ctacaacgag	cgcgagctgc	3120
tgaacagcat	gggcatcagc	cagcccaccg	tcgtattcgt	gagcaagaaa	gggctgcaaa	3180
agatcctcaa	cgtgcaaaaag	aagctaccga	tcatacaaaa	gatcatcatc	atggatagca	3240
agaccgacta	ccagggtctc	caaagcatgt	acaccttcgt	gacttcccat	ttgccacccg	3300
gcttcaacga	gtacgacttc	gtgcccagag	gcttcgaccg	ggacaaaacc	atcgccctga	3360
tcataaacag	tagtggcagt	accggattgc	ccaagggcgt	agccctaccg	caccgcaccg	3420
cttgtgtccg	attcagtcac	gcccgcgacc	ccatcttcgg	caaccagatc	atccccgaca	3480
ccgctatcct	cagcgtgggt	ccatttcacc	acggcttcgg	catgttcacc	acgctggggt	3540
acttgatctg	cggctttcgg	gtcgtgctca	tgtaccgctt	cgaggaggag	ctattcttgc	3600
gcagcttgca	agactataag	attcaatctg	ccctgctggg	gcccacacta	tttagcttct	3660
tcgctaagag	cactctcatc	gacaagtacg	acctaagcaa	cttgacagag	atcgccagcg	3720
gcggggcgcc	gctcagcaag	gaggtaggtg	aggccgtggc	caaacgcttc	cacctaccag	3780
gcatccgcca	gggctacggc	ctgacagaaa	caaccagcgc	cattctgatc	acccccgaag	3840
gggacgacaa	gcctggcgca	gtaggcaagg	tggtgccctt	cttcgaggct	aaggtggtgg	3900
acttgacac	cggtaagaca	ctgggtgtga	accagcgcg	cgagctgtgc	gtccgtggcc	3960
ccatgatcat	gagcggctac	gttaacaacc	ccgaggctac	aaacgctctc	atcgacaagg	4020
acggctgggt	gcacagcggc	gacatcgctt	actgggacga	ggacgagcac	ttcttcatcg	4080
tggaaccggct	gaagagcctg	atcaaataca	agggtacca	ggtagcccca	gccgaactgg	4140
agagcatcct	gctgcaacac	cccaacatct	tcgacgccgg	ggtcgccggc	ctgcccagcg	4200
acgatgccgg	cgagctgccc	gccgcagtcg	tcgtgctgga	acacggtaaa	acctgaccg	4260
agaaggagat	cgtggactat	gtggccagcc	aggttacaac	cgccagaag	ctgcgcgggtg	4320
gtgttgtgtt	cgtggacgag	gtgcctaaag	gactgaccgg	caagttggac	gcccgcgaaga	4380
tccgcgagat	tctcattaag	gccaagaagg	gcggcaagat	cgccgtgtaa	aataaaagat	4440
ctttattttc	attagatctg	tgtgttggtt	ttttgtgtga	ggaaccccta	gtgatggagt	4500
tggccactcc	ctctctgcgc	gctcgctcgc	tcactgaggc	cgggcgacca	aaggtcgccc	4560

ES 2 862 526 T3

gacgcccggg	ctttgcccg	gcggcctcag	tgagcgagcg	agcgcgagc	tgctgcagg	4620
ggcgctgat	gcgggtat	ctccttacgc	atctgtgcgg	tatttcacac	cgcatacgtc	4680
aaagcaacca	tagtacgcgc	cctgtagcgg	cgcatataagc	gcggcggggtg	tggtgggttac	4740
gcgcagcgtg	accgctacac	ttgccagcgc	cctagcgccc	gctcctttcg	ctttcttccc	4800
ttcctttctc	gccacgttcg	ccggctttcc	ccgtcaagct	ctaaatcggg	ggctcccttt	4860
agggttccga	tttagtgctt	tacggcacct	cgaccccaaa	aaacttgatt	tggtgatgg	4920
ttcacgtagt	gggccatcgc	cctgatagac	ggtttttcgc	cctttgacgt	tggagtccac	4980
gttctttaat	agtggactct	gtttccaaac	tgaacaaca	ctcaacccta	tctcgggcta	5040
ttcttttgat	ttataaggga	ttttgccgat	ttcggcctat	tggttaaaaa	atgagctgat	5100
ttaacaaaaa	tttaacgcga	attttaacaa	aatattaacg	tttacaattt	tatggtgcac	5160
tctcagtaca	atctgctctg	atgccgcata	gttaagccag	ccccgacacc	cgccaacacc	5220
cgctgacgcg	ccctgacggg	cttgtctgct	cccgccatcc	gcttacagac	aagctgtgac	5280
cgtctccggg	agctgcatgt	gtcagagggt	ttcacgcgtc	tcaccgaaac	gcgcgagacg	5340
aaagggcctc	gtgatacgcc	tatttttata	ggtaaatgtc	atgataataa	tggtttctta	5400
gacgtcagg	ggcacttttc	ggggaaatgt	gcgcgggaacc	cctatttggt	tatttttcta	5460
aatacattca	aatatgtatc	cgctcatgag	acaataaccc	tgataaatgc	ttcaataata	5520
ttgaaaaagg	aagagtatga	gtattcaaca	tttccgtgtc	gcccttattc	ccttttttgc	5580
ggcattttgc	cttcctgttt	ttgctcacc	agaaacgctg	gtgaaagtaa	aagatgctga	5640
agatcagttg	ggtgcacgag	tggtttacat	cgaactggat	ctcaacagcg	gtaagatcct	5700
tgagagtttt	cgccccgaag	aacgttttcc	aatgatgagc	acttttaaa	ttctgctatg	5760
tggcgcggta	ttatcccgta	ttgacgccgg	gcaagagcaa	ctcggtcgcc	gcatacacta	5820
ttctcagaat	gacttggttg	agtactcacc	agtcacagaa	aagcatctta	cggatggcat	5880
gacagtaaga	gaattatgca	gtgctgccat	aacctgaggt	gataaactg	cggccaactt	5940
acttctgaca	acgatcggag	gaccgaagga	gctaaccgct	tttttgacac	acatggggga	6000
tcatgtaaact	cgccttgatc	gttggggaacc	ggagctgaat	gaagccatac	caaacgacga	6060
gcgtgacacc	acgatgcctg	tagcaatggc	aacaacgttg	cgcaaactat	taactggcga	6120
actacttact	ctagcttccc	ggcaacaatt	aatagactgg	atggaggcgg	ataaagttgc	6180
aggaccactt	ctgcgctcgg	cccttcgggc	tggtgtggtt	attgctgata	aatctggagc	6240
cggtgagcgt	gggtctcgcg	gtatcattgc	agcactgggg	ccagatggta	agccctccc	6300
tatcgtagtt	atctacacga	cggggagtc	ggcaactatg	gatgaacgaa	atagacagat	6360
cgctgagata	ggtgcctcac	tgattaagca	ttggttaactg	tcagaccaag	tttactcata	6420
tatacttttag	attgatttaa	aacttcattt	ttaatttaaa	aggatctagg	tgaagatcct	6480

ttttgataat ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga	6540
ccccgtagaa aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg	6600
cttgcaaaca aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc	6660
aactcttttt ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct	6720
agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc	6780
tctgctaate ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt	6840
ggactcaaga cgatagttac cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg	6900
cacacagccc agcttgagc gaacgacctc caccgaactg agatacctac agcgtgagct	6960
atgagaaagc gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag	7020
ggtcggaaca ggagagcgca cgagggagct tccaggggga aacgcctggt atctttatag	7080
tcctgtcggg tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg	7140
gcggagccta tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ccttttgctg	7200
gccttttgct cacatgt	7217

<210> 46

<211> 49

5 <212> ADN

<213> Secuencia artificial

<220>

10 <223> Señal de poliadenilación sintética

<400> 46

aataaaagat ctttatttc attagatctg tgtgttggtt tttgtgtg 49

REIVINDICACIONES

1. Un elemento regulador de ácido nucleico para potenciar la expresión génica específica de músculo, preferiblemente para potenciar la expresión génica específica de músculo esquelético, que comprende la secuencia de SEQ ID NO:10 y que tiene una longitud máxima de 550 nucleótidos, o que consiste en un fragmento funcional de dicha secuencia que tiene una longitud máxima de 300 nucleótidos y que comprende la secuencia de SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40 o SEQ ID NO:41.
2. El elemento regulador de ácido nucleico según la reivindicación 1, en el que dicho fragmento funcional consiste en la secuencia de SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40 o SEQ ID NO:41.
3. Uso *in vitro* o *ex vivo* de un elemento regulador de ácido nucleico según la reivindicación 1 o 2, en un casete de expresión de ácido nucleico, o en un vector.
4. Un elemento regulador de ácido nucleico según la reivindicación 1 o 2 para su uso en medicina, en el que dicho elemento regulador de ácido nucleico está en un casete de expresión de ácido nucleico, o en un vector.
5. Un casete de expresión de ácido nucleico que comprende al menos un elemento regulador de ácido nucleico, operativamente unido a un promotor y un transgén, en el que dicho elemento regulador de ácido nucleico consiste en el elemento regulador de ácido nucleico tal como se define en la reivindicación 1 o 2.
6. El casete de expresión de ácido nucleico según la reivindicación 5, en el que el promotor es un promotor específico de músculo.
7. El casete de expresión de ácido nucleico según la reivindicación 5 o 6, en el que el promotor es el promotor del gen de desmina (DES), el promotor de SPC5-12, el promotor de creatina cinasa muscular (MCK), el promotor de dMCK o el promotor de tMCK.
8. El casete de expresión de ácido nucleico según una cualquiera de reivindicaciones 5 o 6, en el que el transgén codifica para una proteína terapéutica, una proteína inmunogénica o una proteína estructural, preferiblemente distrofina o un sarcoglicano.
9. El casete de expresión de ácido nucleico según la reivindicación 5, en el que el transgén codifica para distrofina, un sarcoglicano o folistatina.
10. Un vector que comprende al menos un elemento regulador de ácido nucleico, en el que dicho elemento regulador de ácido nucleico consiste en el elemento regulador de ácido nucleico tal como se define en la reivindicación 1 o 2, o el casete de expresión de ácido nucleico según una cualquiera de las reivindicaciones 5 a 9.
11. El vector según la reivindicación 10, que es un vector viral, preferiblemente un vector viral adenoasociado.
12. Una composición farmacéutica que comprende el casete de expresión de ácido nucleico según una cualquiera de las reivindicaciones 5 a 9, o el vector según una cualquiera de las reivindicaciones 10 u 11, y un portador farmacéuticamente aceptable.
13. El elemento regulador de ácido nucleico según la reivindicación 1 o 2, casete de expresión de ácido nucleico según una cualquiera de las reivindicaciones 5 a 9, vector según una cualquiera de las reivindicaciones 10 u 11, o composición farmacéutica según la reivindicación 12 para su uso en terapia génica.
14. El elemento regulador de ácido nucleico según la reivindicación 1 o 2, casete de expresión de ácido nucleico según una cualquiera de las reivindicaciones 5 a 9, el vector según una cualquiera de las reivindicaciones 10 u 11, o composición farmacéutica según la reivindicación 12 para su uso en terapia de vacunación, preferiblemente vacunación profiláctica.
15. Un método *in vitro* o *ex vivo* para expresar un producto transgénico en células musculares que comprende:
 - introducir el casete de expresión de ácido nucleico según una cualquiera de las reivindicaciones 5 a 9, o el vector según una cualquiera de las reivindicaciones 10 u 11 en las células musculares;
 - expresar el producto transgénico en las células musculares.

FIG. 1

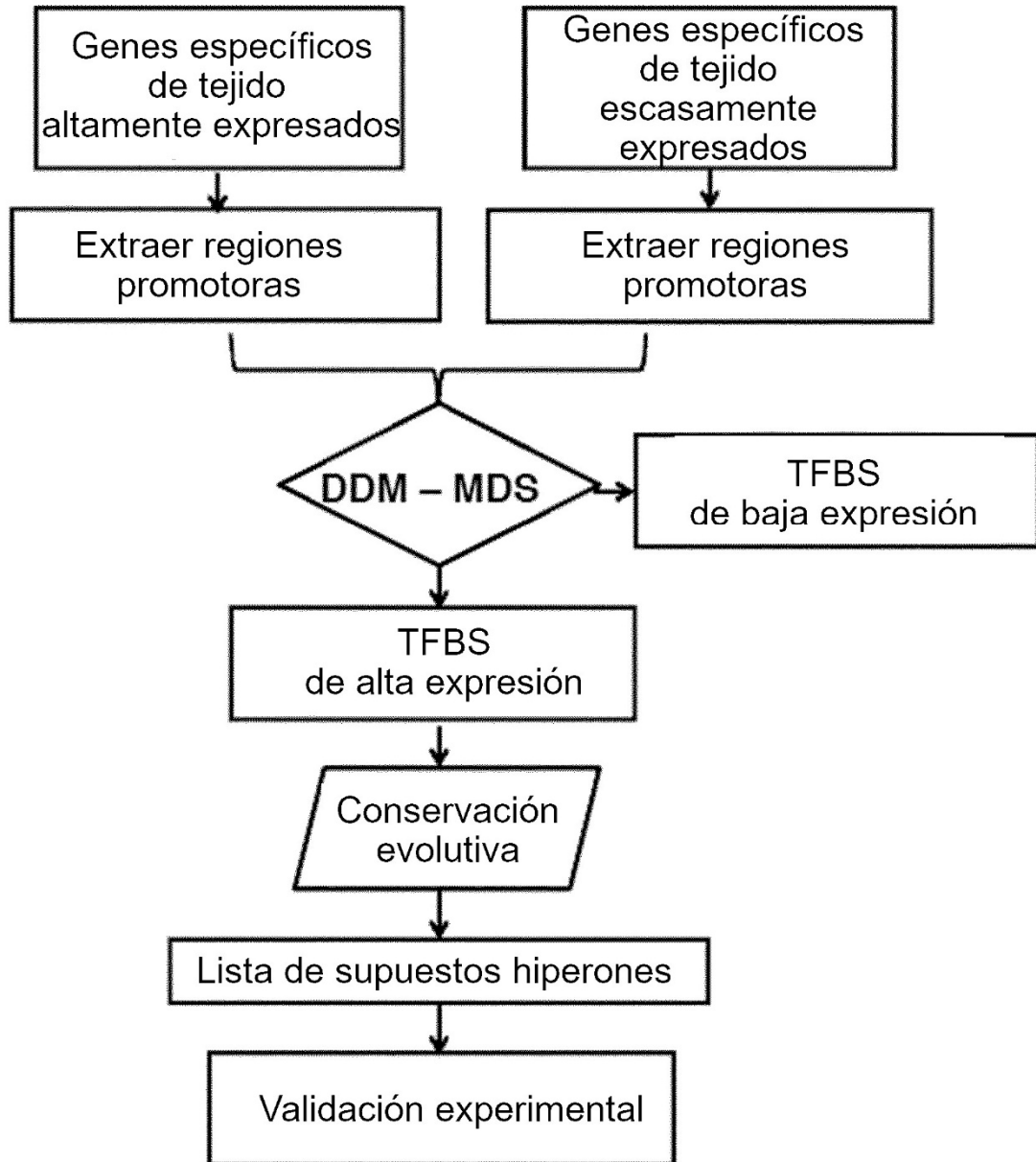


FIG. 2

A

SEQ ID NO: 1

GTTCTCCTCTATAAAATACCCGCTCTGGTATTTGGGGTTGGCAGCTGTTGCTGCCAGGGAGA
TGGTTGGGTTGACATGCGGCTCCTGACAAAACACAAACCCCTGGTGTGTGTGGGCGTGGGT
GGTGTGAGTAGGGGGATGAATCAGGGAGGGGGCGGGGACCCAGGGGGCAGGAGCCACACA
AAGTCTGTGCGGGGGTGGGAGCGCACATAGCAATTGGAACTGAAAGCTTATCAGACCCTT
TCTGGAAATCAGCCCCTGTTTATAAACTTGAGGCCCCACCCTCGACAGTACCGGGGAGGA
AGAGGGCCTGCACTAGTCCAGAGGGAACTGAGGCTCAGGGCTAGCTCGCCCATAGACATA
CATGGCAGGCAGGCT

B

SEQ ID NO: 2

CGTCTATAAATTCCAGGGAAGGTCTCTGATTGGCCCTGCTCATTCCCAGGCCCATTCTTG
ACCCAGTCACTGAAGTCAGGGAGATGCAGTAATAAGACTGGCTGGAATCAGGGTCTTTAGG
GGTGGAGGGATGGGGAGGAGGCACAGCATGTCATCAAATAAGGAAATTGCAAAAGAAAGC
TTGCAGGCTACTTTGAATGACAATGAGAAAGACGGTGCTGCCTGAGTGTGTTAAGGATCCA
CATGGTCTCCAAAATCCTCCAGGAGCATAACAGTCTAGTCTGGGAGATGAGACACAAAAATA
ACCAGAACACAACAGCTTGCCTGACTCGAGGGCTGGATAAGAATATCTGGAACCTCCCCCA
TCTATTTTCAAGCTTGTCTCTTGGATGAAAATTAGACACTTAATGGGAAAGGGCTTTGAA
AAGAGTGC

C

SEQ ID NO: 3

AAAAAATAAAAAATAAAAAATAAAAAATAAAGTGGATGGCTTAGAGTATCTTTTCTGTT
TAGACCTGACTAAAGCTTAGACATAATTGTTAGTTTAGGCTCTCAGGGTAAATTTATTAC
TGTAATCCAAAAATCCCTTCTTCTCTTTTTTTTTTTTTTTTTTTGGTCTTGAATTAAAT
GCTGTCACCTCCTTCTTGAAAGGAGAACTATTAGTCAGATTTGAAAATCCTCTTTATCAC
CCAGGAAAATCATTTTTATGGACACTTTGTCTTTCTGTAGTCTGACTTAGAAGCAGCCTGT
TTTTGATAGGTTGAAGTTTTATCTTGAACACAAACCTGTTTGTGTGTCCCCTACTCCCC
AGTTTGATGTGCCAGGCACTTTGTTCTCAAGCCCAGCAGCTGTTGTGGGATGAGGGGACAT
TTTGCATGCTTAGCCAGCAGCTGCCAGAAACATTTCTAATCTGGTTTTTGGCAGGAAATAGG
GCACAAGTGGAAGCCAAGTTAAAAGAAGCTGGAAAAATAAACAGAATAACTTTAGATGTCA
CT

FIG. 2

D

SEQ ID NO: 4

TGCACTGTAGACTATTTATGTTTGCATATTTTTCAGGATTGTGTACAGTAAACCTTAAGTG
AATACAAAGGATCCAATTGTCCTGTAAGACTCACTTCAATTACAGATTGTGCTCAGTATTA
AACTTTGCTAGTACTTTTCAGGATGCAGTCAATCAAGAGGCAGGGAAAGGTCTGTCAGCATC
CACAGCCTCCTTTTCAAATTGGACACTTAGTCTCTGGTCCGAATAATAGGTGCCCCCAGTG
TCACCTCACACAATGGGGTATACACAGTTTTTAAAAATTTATTTTCAGCTGGGCACAGTAG
CTCACACCTAAAATCCCAGCACTTTGGGAGGCTGAGGCAGGTGGATCACCTGAGGTCAGGA
GTTCAAGACCAGCCTGGCCAACGTGGCAAAATCCTGTCTCTACTAAAAATACAAAAATTAA
CCG

E

SEQ ID NO: 5

CAGTTTACTCACCAGGGATTTCAGAGGCAGCACTGCTGAACCCTGAGCCCTTGGCACATCAG
GTTGGCTGTCAGAAGTCGGCCTTTGTACATACACAGTTCCCTTGTGAGGCCCAGCTGCGTG
TCCTAGGAGCGGGGCCTCTCTCCACAGCAGAGCTCAGCCTCTCAAGTGTATGGACAGCACG
GGTGCTGATGGGTGGATTTAGCCATGAGTTGAAGGTGGCTTGGGGAGAATGAGAGTTCTA
GAGATAGGGAGAAGGGGTTGCCAATAGGAGAGTGGAATTCCTGAGCACCTCGTCACAGGCA
GCCGACAGAACATGAGCCGCAGGGCCCAGGCTATTTATACCTCGCCTGTCACTATCAGGGT
CCCCACAGCTCCCCCACCTCCAGCCACACACAGCAGGTCCTTTTGCTCTTTCTGGTCCCT
TCTCTACTCCTCCCCCTCCCTACCTAA

F

SEQ ID NO: 6

CCCTTTCCCTGGCAGGATCTGCCCTGTGGCCCAAATGGGCATGTTGCCCAGGGGGCTCCCT
GGCACTATGGGGGAAGAGTCTCTCCTTCCCCTCTTATCATCTCAGTTGAGTCAGACTTGGG
GGAGGGGGATACACAGTGTGAGTCACTGGGTACCCTTTTCTGAGCTCAGCTTCATACCGA
GGCGATGAGGCCAAACGGGCTGGTGACAGGGACACTGAGTCAGGGGCAGGGGCCCCGGTCT
TACTCCTGGGCCTCTGGATTTGGGCCCTACATGAGGCTTTTCTATCTGTAAAGTCAAGCAA
TGGCTGGGAGGCACACACAACCCCCCGCCCCCGCAGGCTTCTCCTTCATTGGCCCCGGGC
AAGGTCCCTGCTTCCTCTCAGGCCGTCTCTGCACAAGCACACACACTTCCCTTCCCTGTCC
ACAGGTGGACAATGCCCTGGGCTAGG

FIG. 3



FIG 4

A

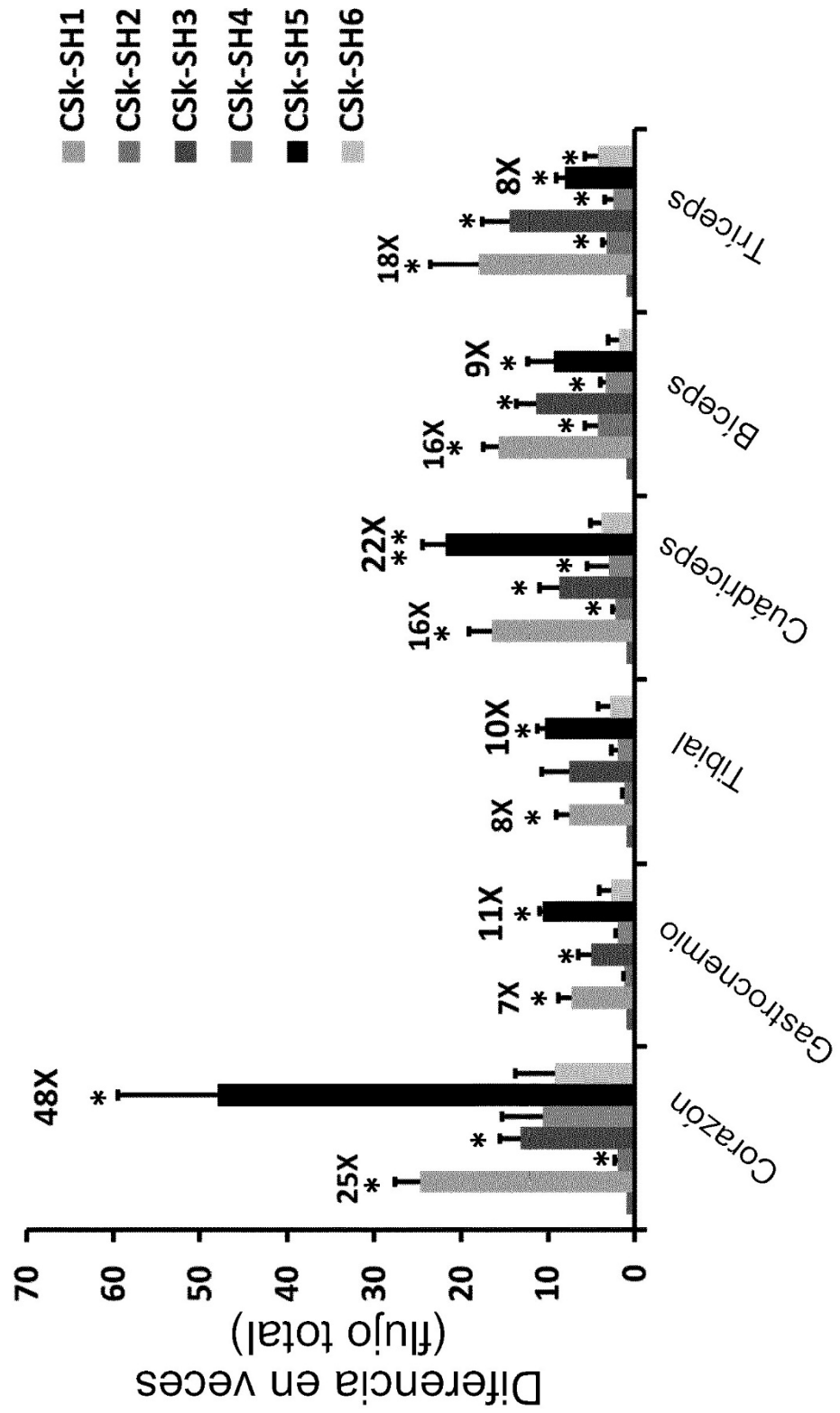


FIG 4

B

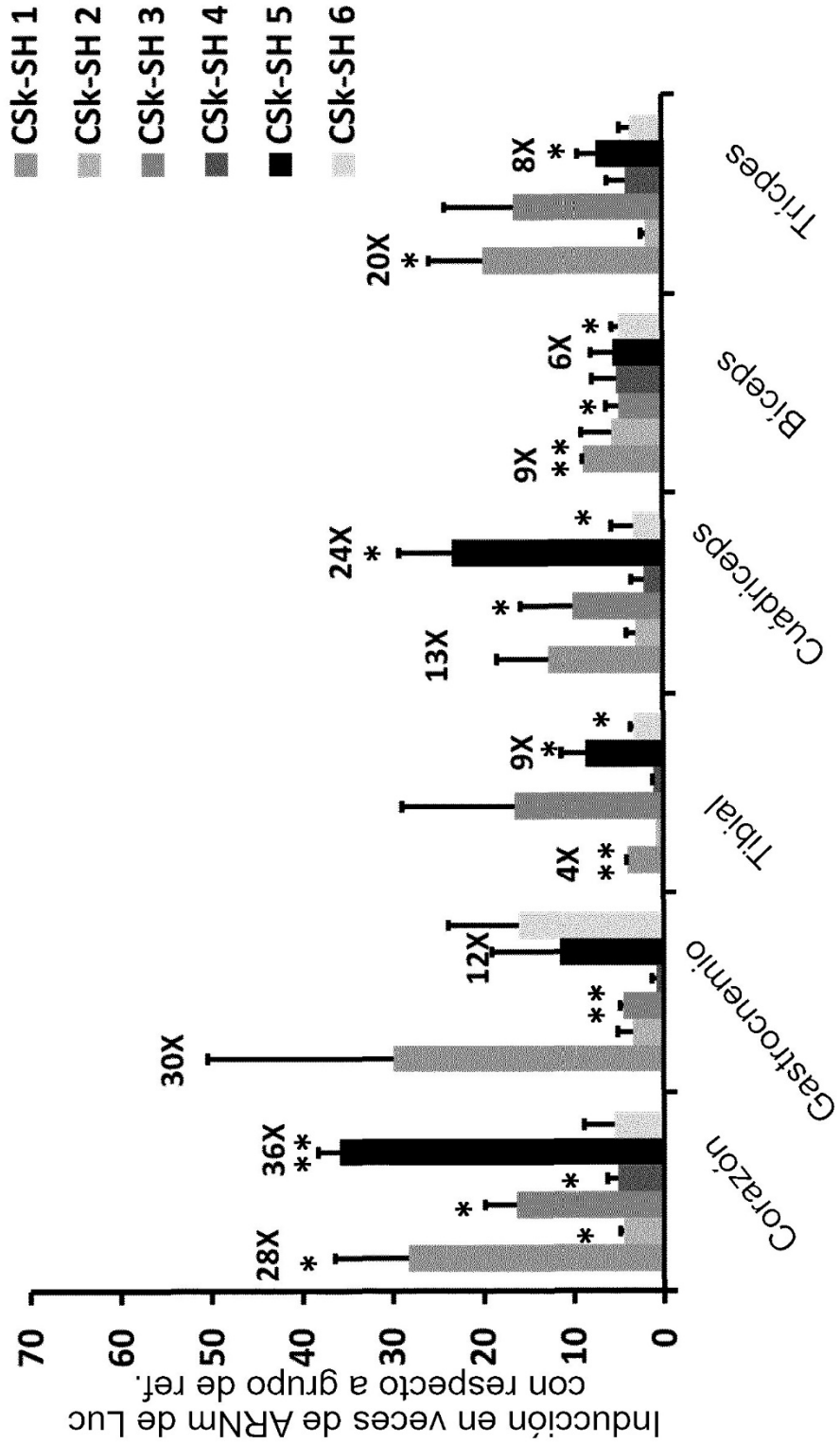


FIG 5

A

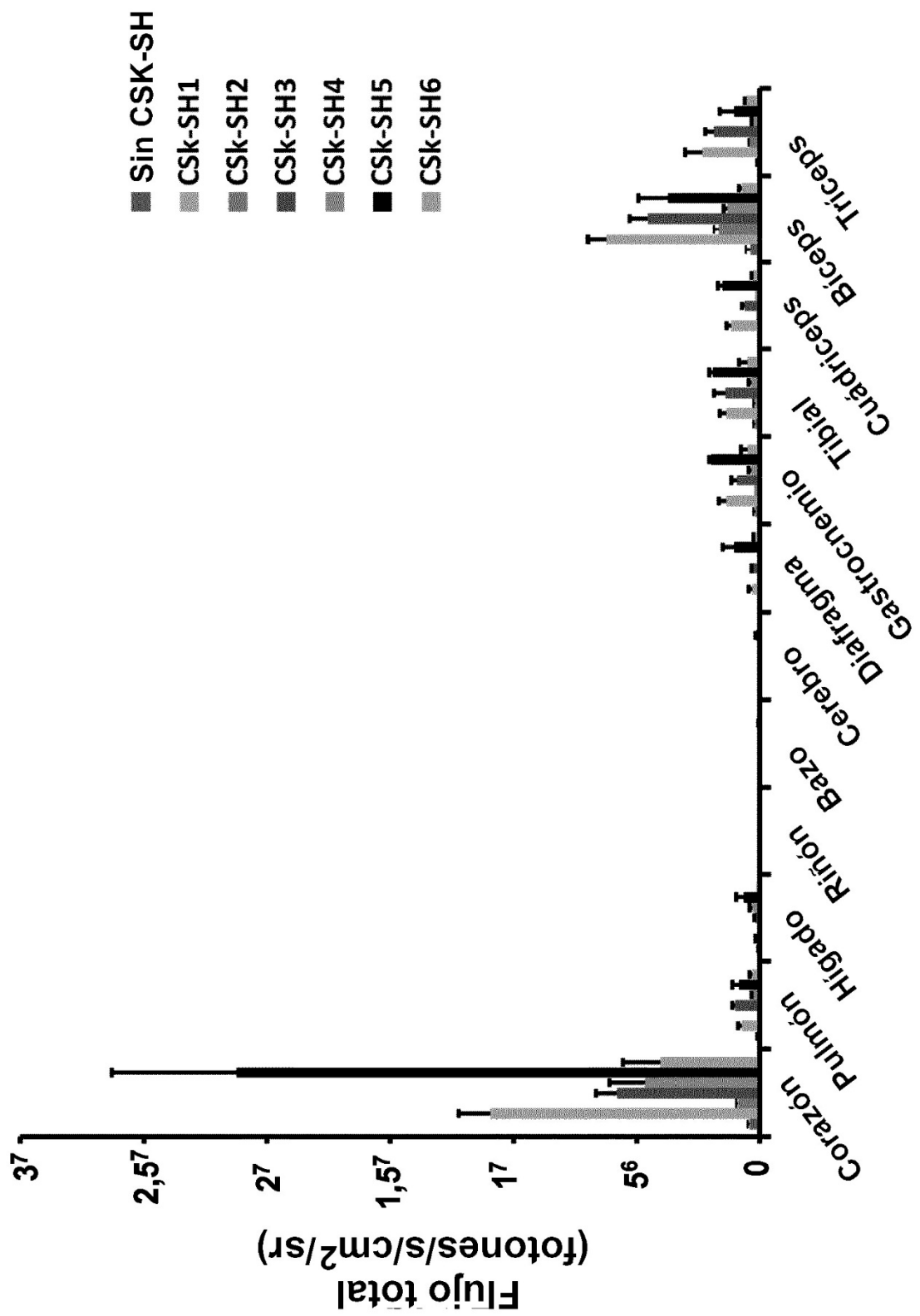


FIG. 5

B

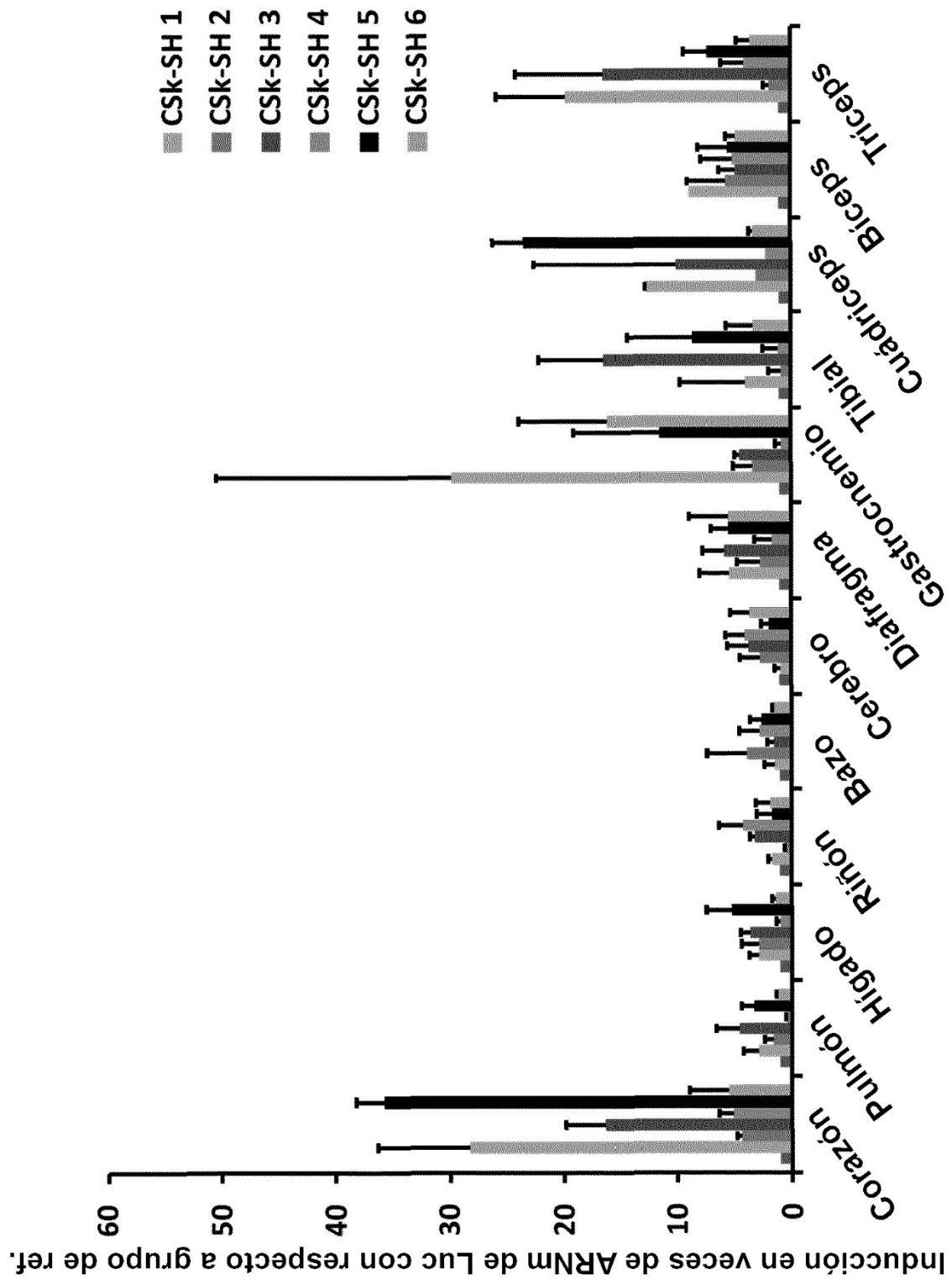


FIG. 6

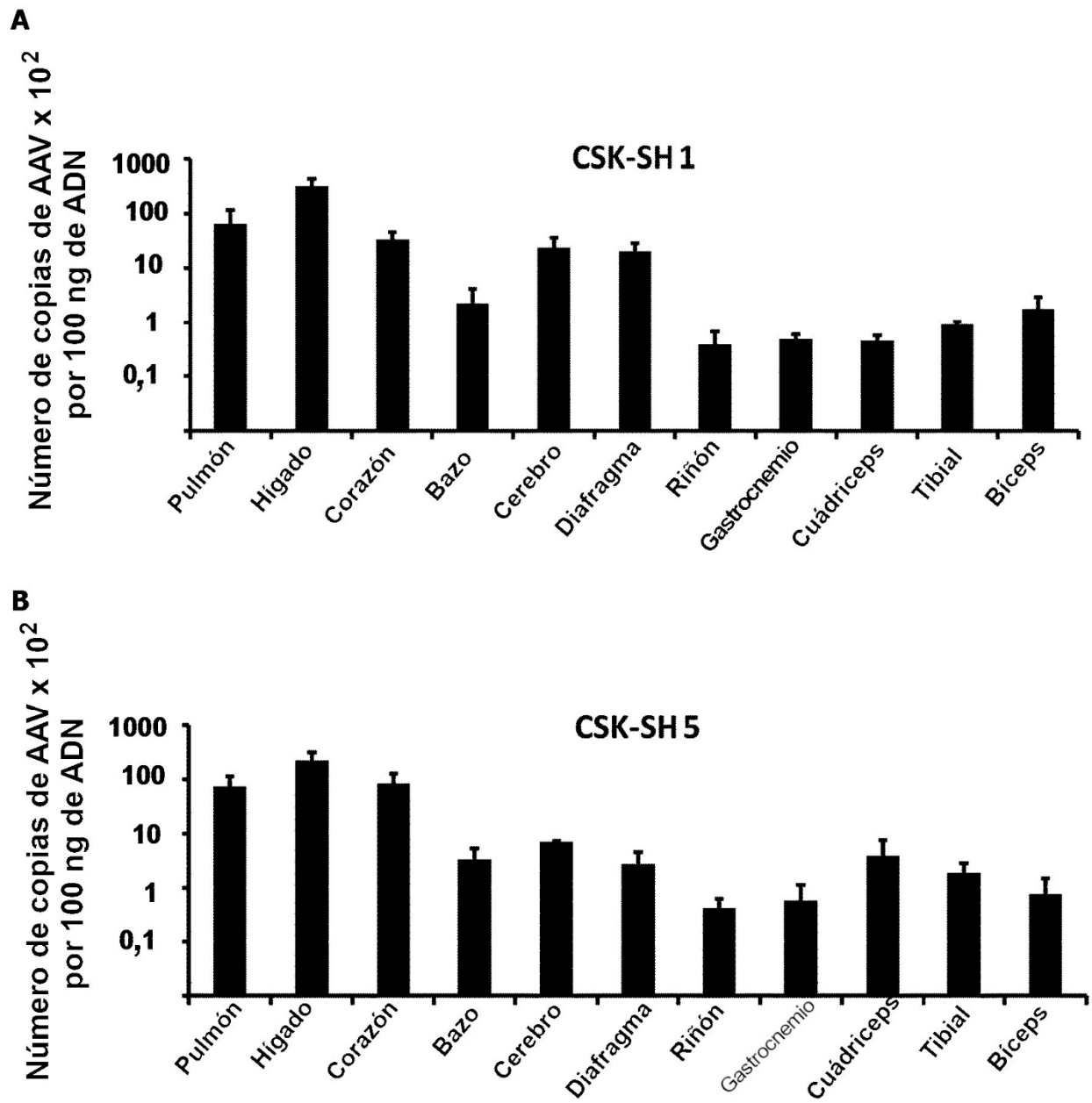


FIG. 7

A

SEQ ID NO: 7

CTCACTCCCCGCCCAGGCAGCAAGGAGCCCACACCCTCATGCCCCCTCAGCTTCAGCCCCCA
CCTCCAGGAGGCCCTACCCACGCTCATGACCTTGCTATTCTGGGCCTTGCTGTCCTGTAGGG
AGATGGACAGGAGACAGCTGGGCTTCCAGGCCACCCAGGCGGGGGGCTAGCCGAGGGAAGC
CTGCTGGCTCTCCTGCTTGCTCTAATTTCTGGGGCTCCCCAAACCTTGGCCTCAGGAGACT
GGGGATAGGACCGGCCTTGAAAGTGGGGGAAGCTTTGGAGAGCCGGGTGCTGGGTTCTTAG
TGAGATGGCCAGTGAAGGCTGTGGTGCCCCGAGGTAAGCAGGGCCTGATCCCCCTCCTAATC
TTCCAGCAGCAACTGGTGCTCTGAGGCTCCCCCTCCCCCAGCCCTGCCAGCCTTCAGGGAC
CTGCCTTCCAAAGATGGGCAGGGGAGGGGACGAGGACACCCACCCACTCCTCAGACCAGC
ATGTCTT

B

SEQ ID NO: 8

CCCTCCAGATGGGTTTCCTGGAATCTAGATTTCCCAGGTTCCAAAGGACACCCGAGTCTCA
TGCCTGGAACCTCAGTGAGACTAATTCACCTCTCCTCTGCCCTAATCTTCATCTCCAGCCAG
AAGCCAACAGATCCCAGGGGACTGGAGCCACAGGGGCTGCACCTGTTTACCGGGTATTTTT
AGGATGGTTGATGAACACATAATACCCACCCTATAGTCAGAGAAAGACAATGCCTGCTATG
TTAATCCTGTGGCTATTATAGTCTGTCTCATCTCATGGGTTGGGGCAGGACACTGACCCTCTC
AGAGGCCAGAGAGAGGCCTCGCAAGCAGGAGGTTAGGGA

C

SEQ ID NO: 9

ATGGAGACAATCCATGAATTCCTGAGATGCTTGGCTGGTATTAGATTTTATGGGCAGCTGC
TTATTCTTAGGGCTCTGCTTCTCCAAAGACACTGAGGAAGTCCAAAGGAAACACCAGCTGG
CGAAGAGCCACCTCCAGGCCCATCTGTCCATCATCAGCCTCCAGGAATGCCAGTGTCCAGA
GGGCACCAGGTCTGCGTCTGTCTCCCTGGGATGTGCCTTGTCTTGGTGGGCATTTGGCAG
TGATCATGCCTCCCTGTCTCCCTCAGAGATCCAACTGTCCCCATTGTGGGGCCCTACCTTC
CAAGGCCGTTTACACCTCCTGCCAAGCTCCGGGGCCTGCCCCAGCCTGCCTCACTGACA
AATGCCAGACCAAGGGGTCCCACGTCAGGCAAGAGGCCTCAGCCTGTGCTCTGACACCCCT
CAG

D

SEQ ID NO: 10

TTCTGAGTCCTCTAAGGTCCCTCACTCCCAACTCAGCCCCATGTCCTGTCAATTCCCCTC
AGTGTCTGATCTCCTTCTCCTCACCTTTCCCATCTCCCGTTTGACCCAAGCTTCCTGAGCT
CTCCTCCCATTCCCCCTTTTTGGAGTCCTCCTCCTCTCCAGAACCCAGTAATAAGTGGGCT
CCTCCCTGGCCTGGACCCCCGTGGTAACCTATAAGGCGAGGCAGCTGCTGTCTGAGGCAG
GGAGGGGCTGGTGTGGGAGGCTAAGGGCAGCTGCTAAGTTTAGGGTGGCTCCTTCTCTCTT
CTTAGAGACAACAGGTGGCTGGGGCCTCAGTGCCCAGAAAAGAAAATGTCTTAGAGGTATC
GGCATGGGCCTGGAGGAGGGGGGACAGGGCAGGGGGAGGCATCTTCCTCAGGACATCGGGT
CCTAGAGG

FIG. 7

E

SEQ ID NO: 11

GACTAGGAATAAATCACATATCCTCAATCCCTGGACAACCTTGTTTACTTCTAGTGTTAGTT
TTTTCTTAAAAAAAATTGAAATCATTCTGAGGCTGGAATACTTTGGACATGCCCAGCAG
TTCCTGGCAGTTCCACAGAAGCATTACCTCATGACTGGAGTGGGTAAAGCATACTGTGGG
CTATGGATAAGACTGACATTAACCACAAGCATGTTTGGCAGCAGACTGGTGCTTTACAAGC
TCCATGTTTCAGCAGGAGCTGCAAAGTGTTCCCTCCAAACCAATATTTGTCAATTCTTGGATT
TATTTAGGAGGTCCTGTTACTCACATGTTTCAATATCAGCAGAAGCCAGTTTCCCTGTGGT
ACCGAAGTGGATCCTGATGAATTTACCCTTGTAAGTAAAAAAAATGATGTTATACCCAAAG
CTTGAAGTACGTAGTGGGGATGCCACTGAAATAATTCAGACATGCTT

F

SEQ ID NO: 12

GTGCTCATAGCTCCACCTTTTGTTCCTAATATGGTCTTTCCAGCTCCCTCCACCCCATCAT
TGTTCTCCTGGGGGAACACAGGGTGAGACGCTTTGATGAACTGACATCACCAGCAAAAAA
ATATCTAGCAACAGCTGAGGCTGATTTTAGACAATGGAAAGTGGGGGAGGGAAGAGGTTCT
CCCTGACCCTGAAACTTTCCACTCATTCTGGGCAGCTCTATGGATGTTTTAAAAGAAGAGG
AAGAGGGGAGGGAAGAACATTGAAATAGAGAAGTGTAAGTTTGGCAATTCTAGGTTGGCAGT
TTGCATCCAGGGGGTCCTGGTTGCCTTTCAGCTTCCCGTTTCACTCTCCCCAGACTGTGT
TGAATGCTGGTCAAACCTCCGTTAGTTGAGTTTTAGCTTTTGATTCTGGTATTCAAGGAGC
TTGGGCACAGGGAAGAGGGGAGGTCATCATGATCCTTAACAATTCTCCCAGATCCCCAGA
TCAAATTGCTGTGCTATTCTGGGAGTCTCCG

G

SEQ ID NO: 13

ATCGTGTGTCAGAGGTTTGTGTCAGCTTCCCAGCAAGGGAACCAGAAAGGAAAAGGAACCG
GTTCCCTCATGCTTCCTAGGGGAATGCATGCATATCTGAAGAGAAGGGAATCTTATATAAGG
CTGTTTAGCTAAGGGCAGCCACCAGCCAGGTGAGCCTTACAGAAGCACAGGGCTGGGTGTC
TGCAGTTCCCTAGCAGATTAACTGGGTACAGTGACTCAGAGCTCCAGCATGCGAGTTCC
AGGTGTGGAAGTACAGATCTGCTTTTGCTCCACTTGGGAGTATTTTTCCTTCT
TAGTGAGCATGGGCAGCCTCCTGGCCAGGGAAGTCTGGCACTGTCTGGGCCTGACAGGGAA
ACCCTG

FIG. 8

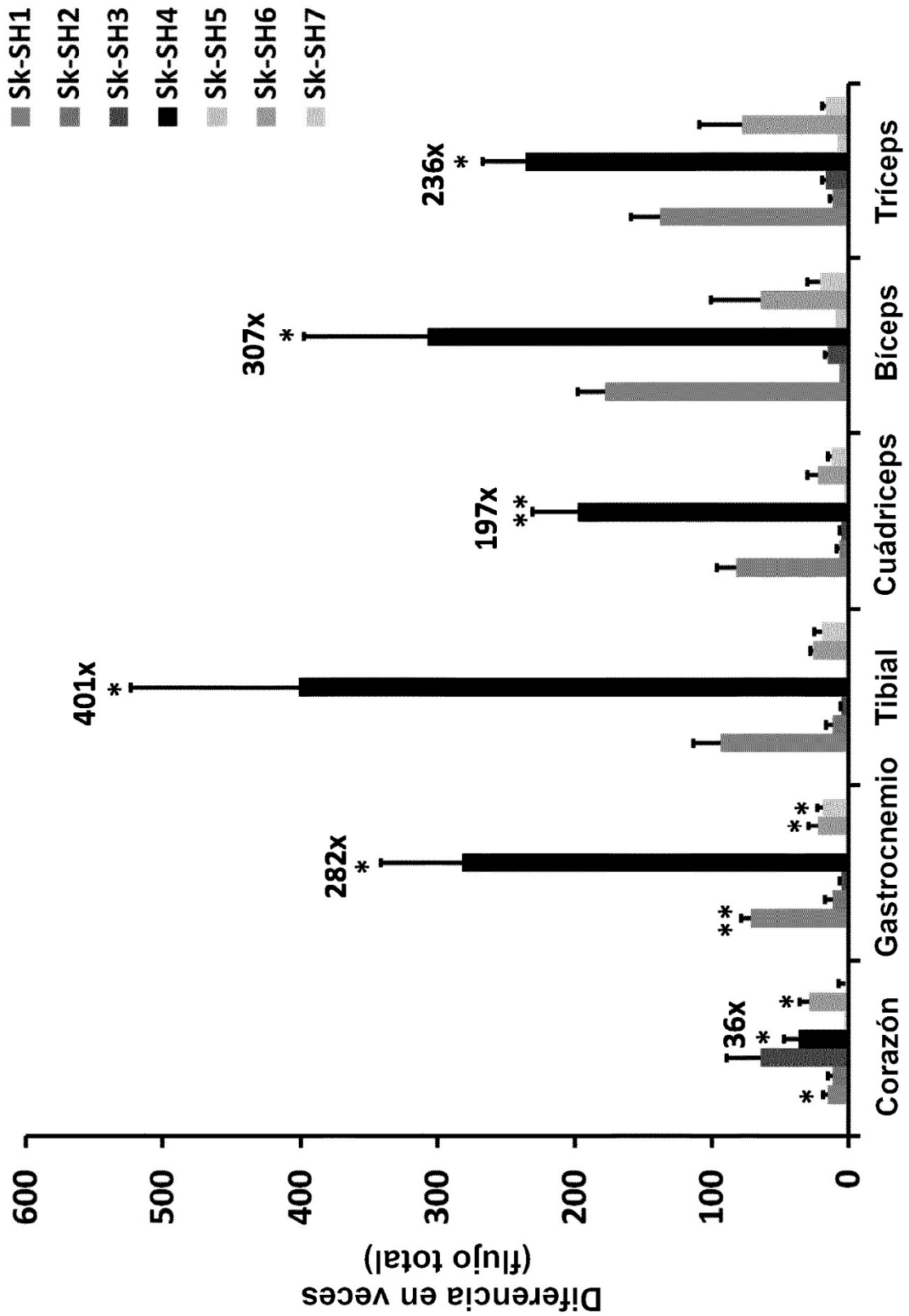


FIG. 9

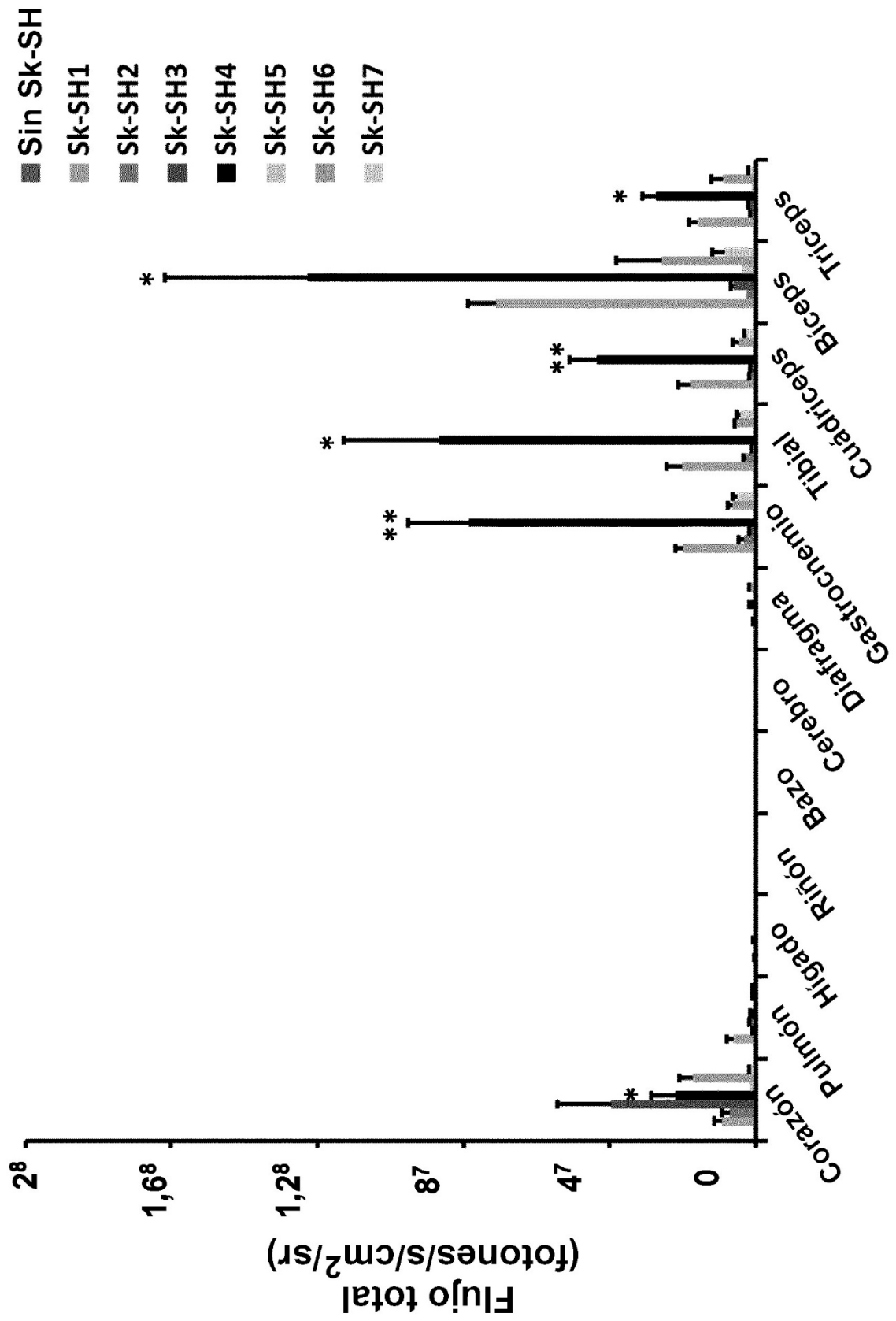
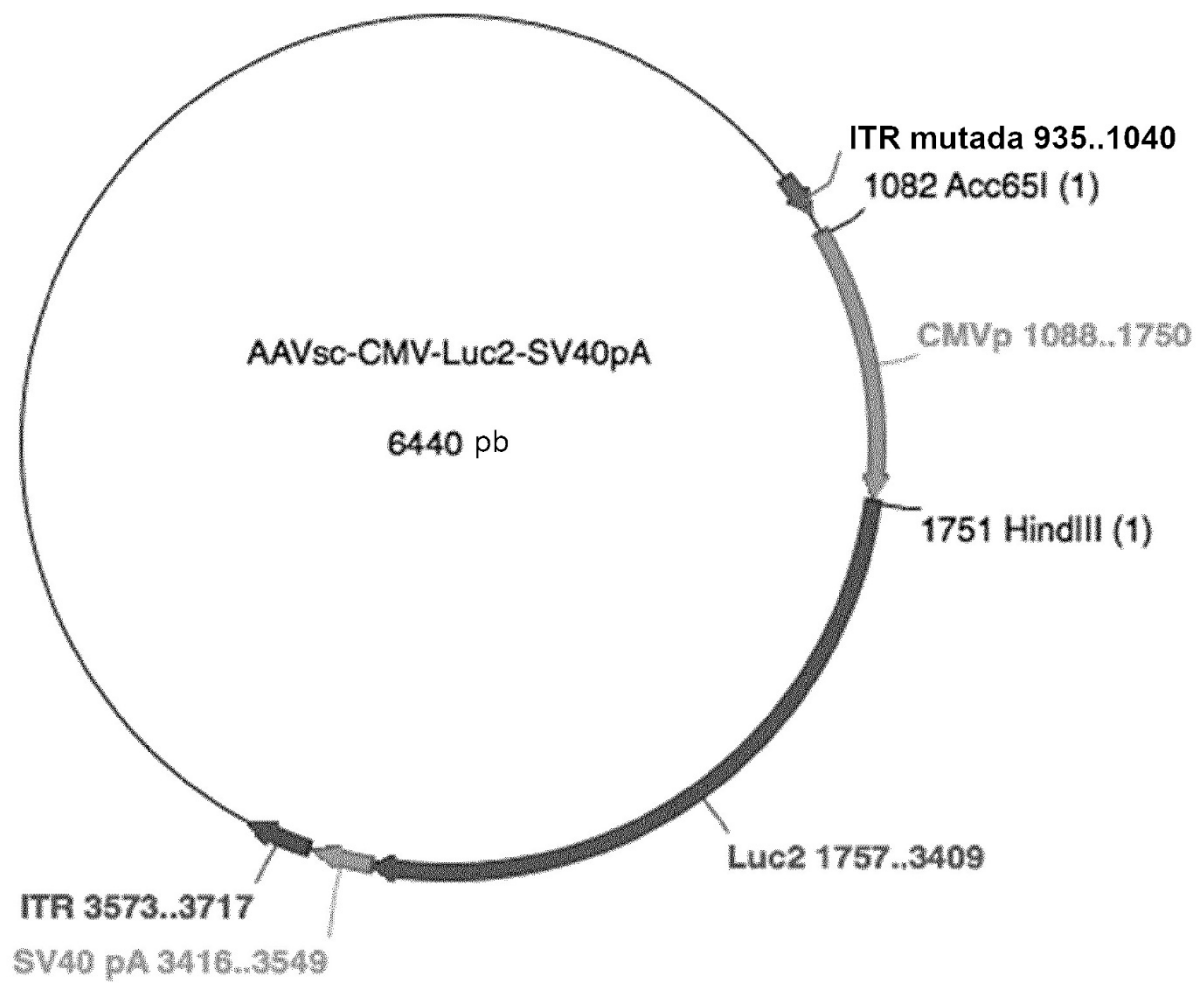


FIG. 10

A.



B

FIG.10

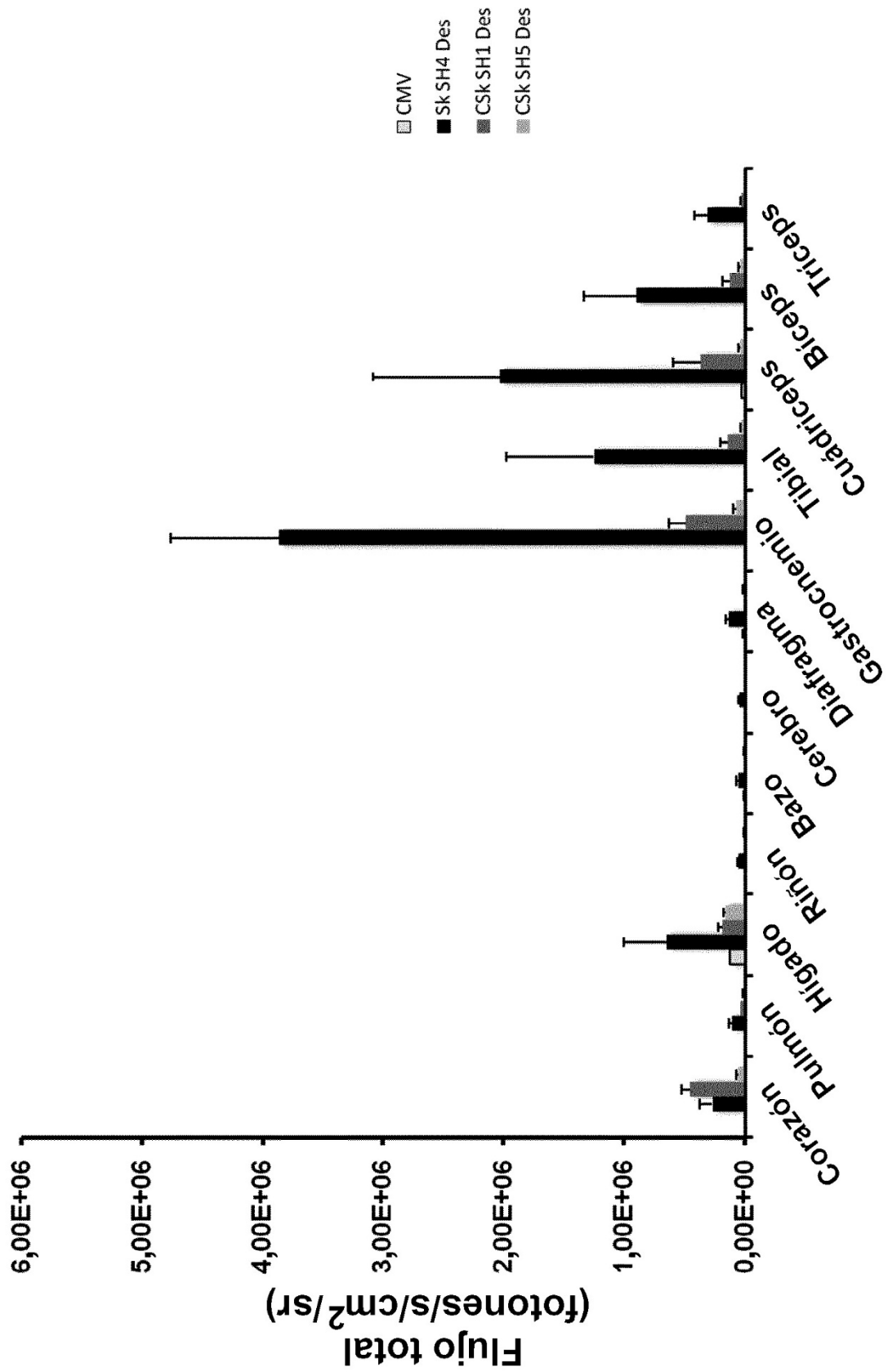
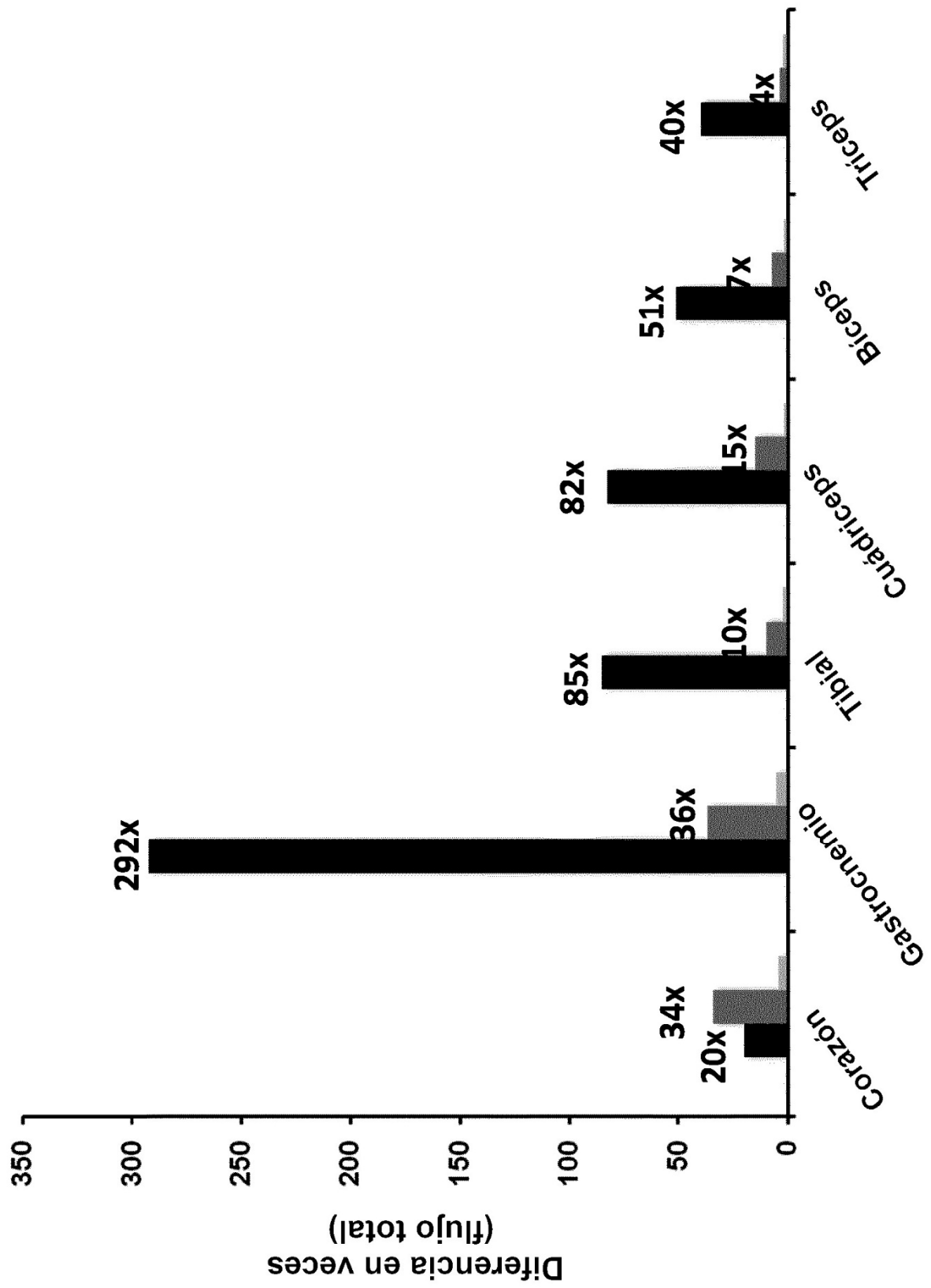


FIG.10

C



A

CSK-SH 1

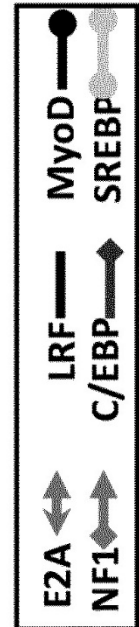
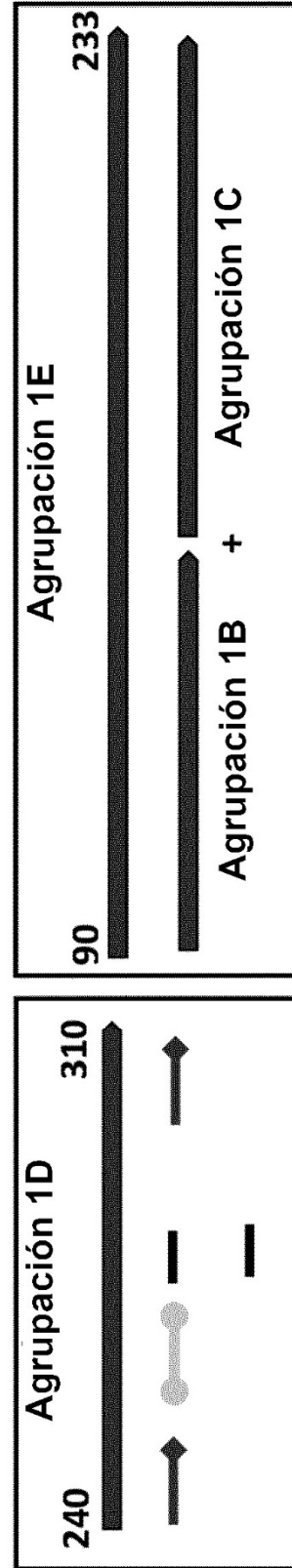
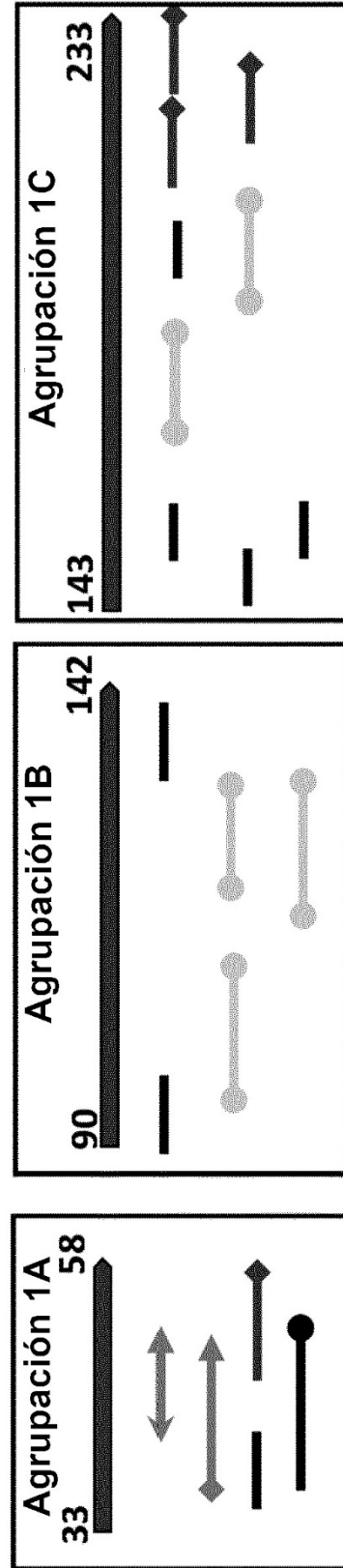
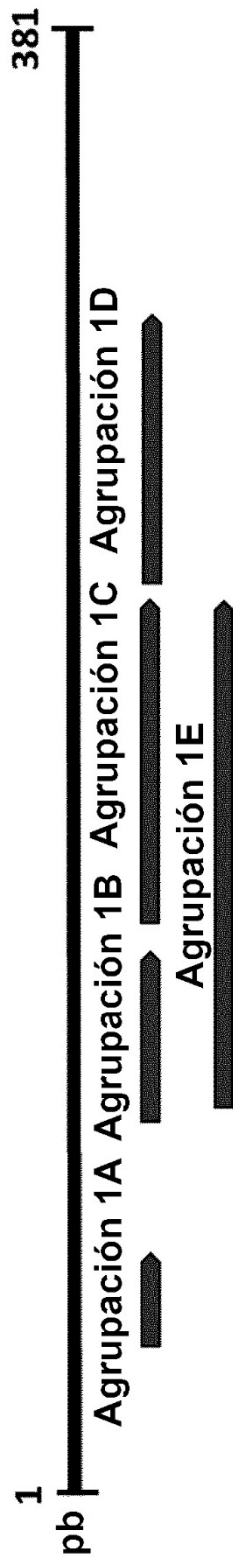
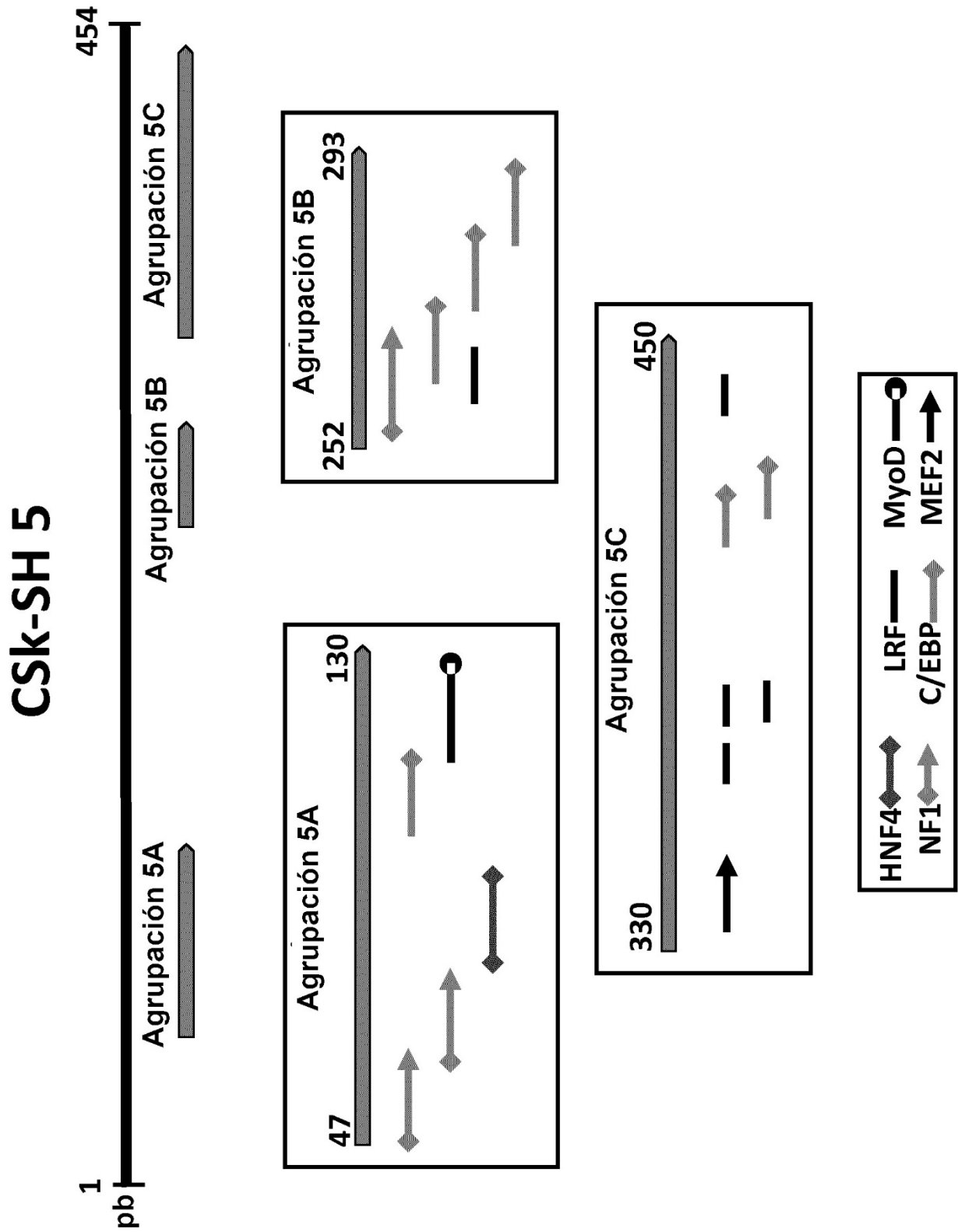


FIG.11

FIG. 11

B



c

Sk-SH 4

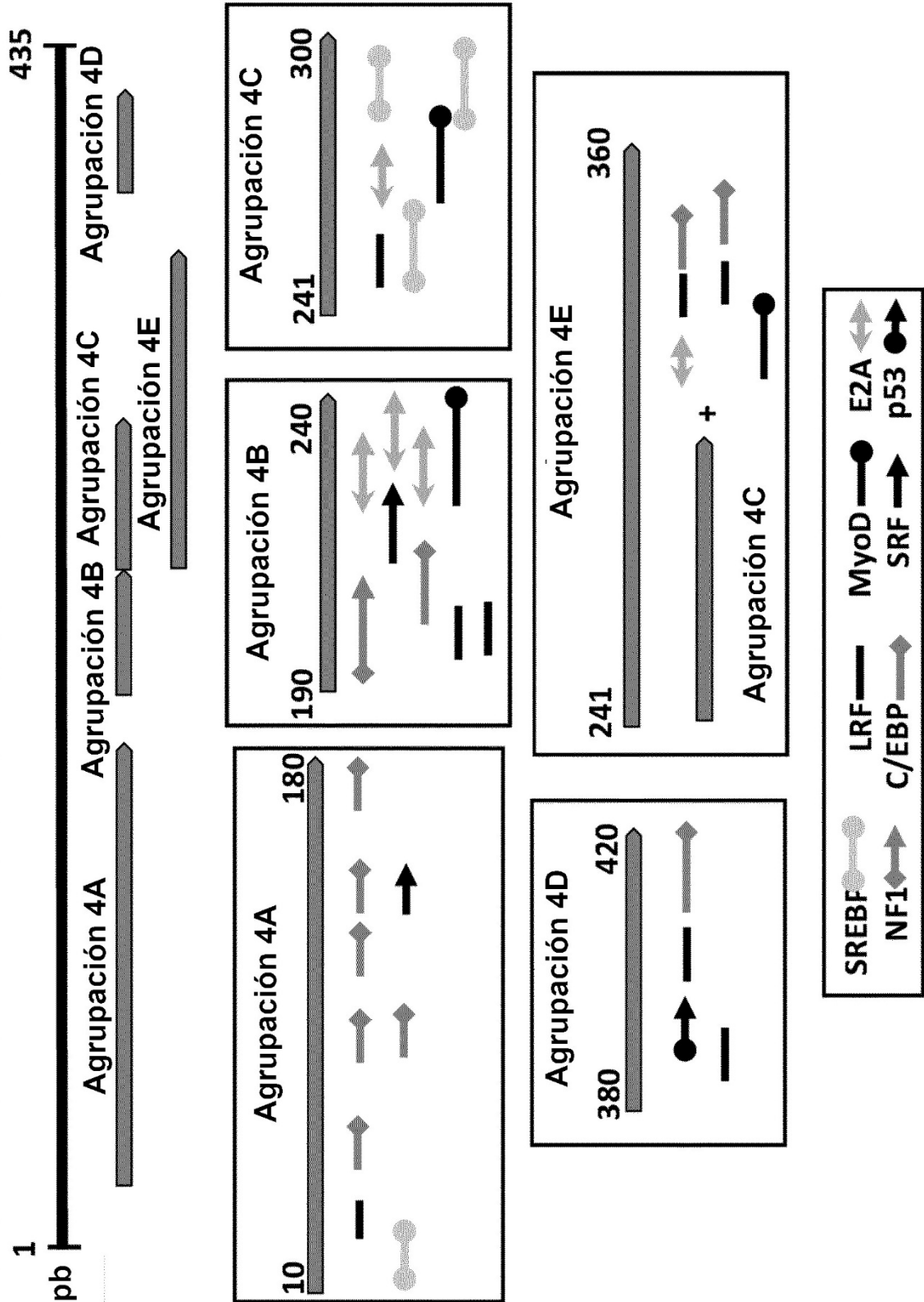


FIG. 11

FIG. 12

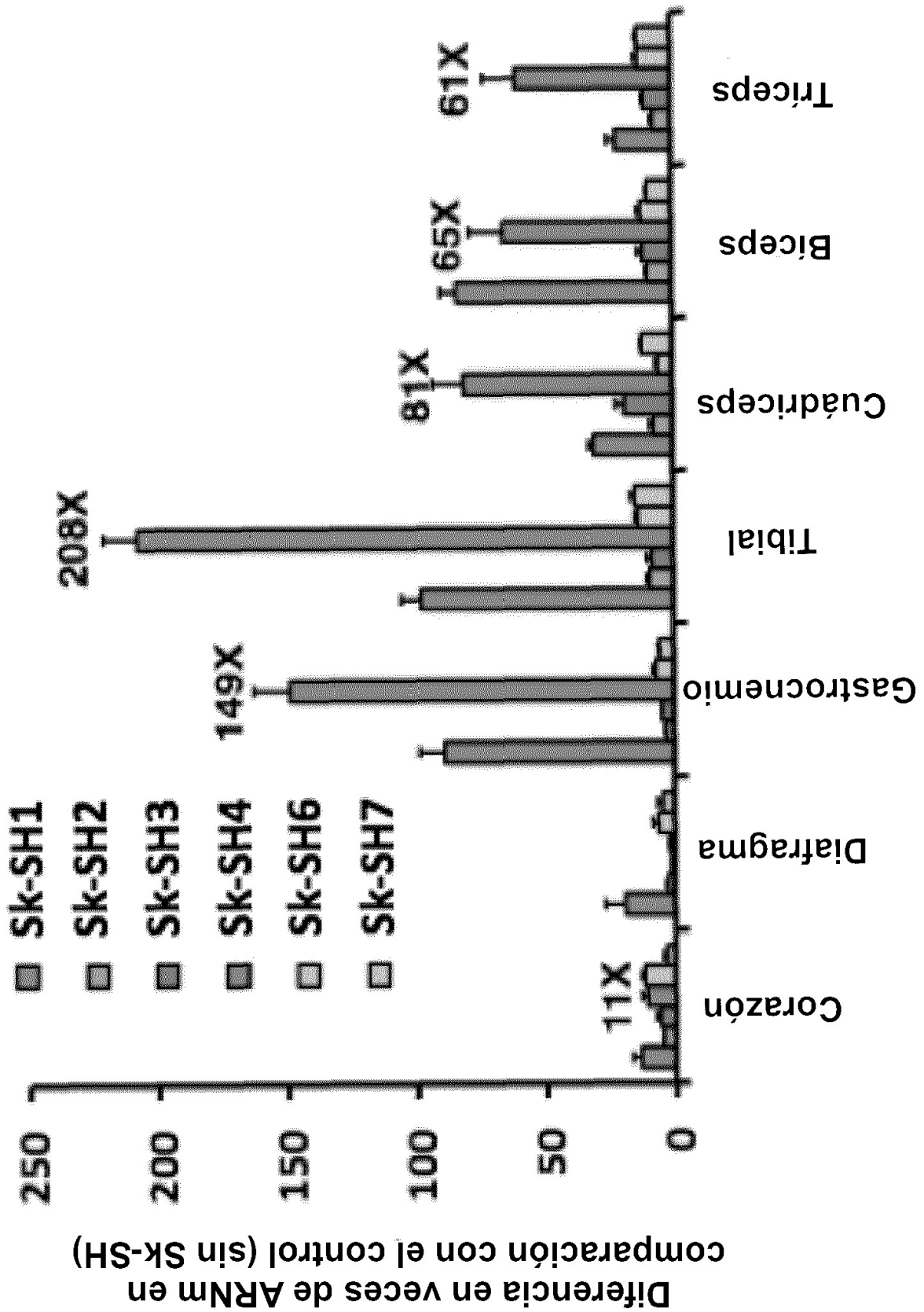
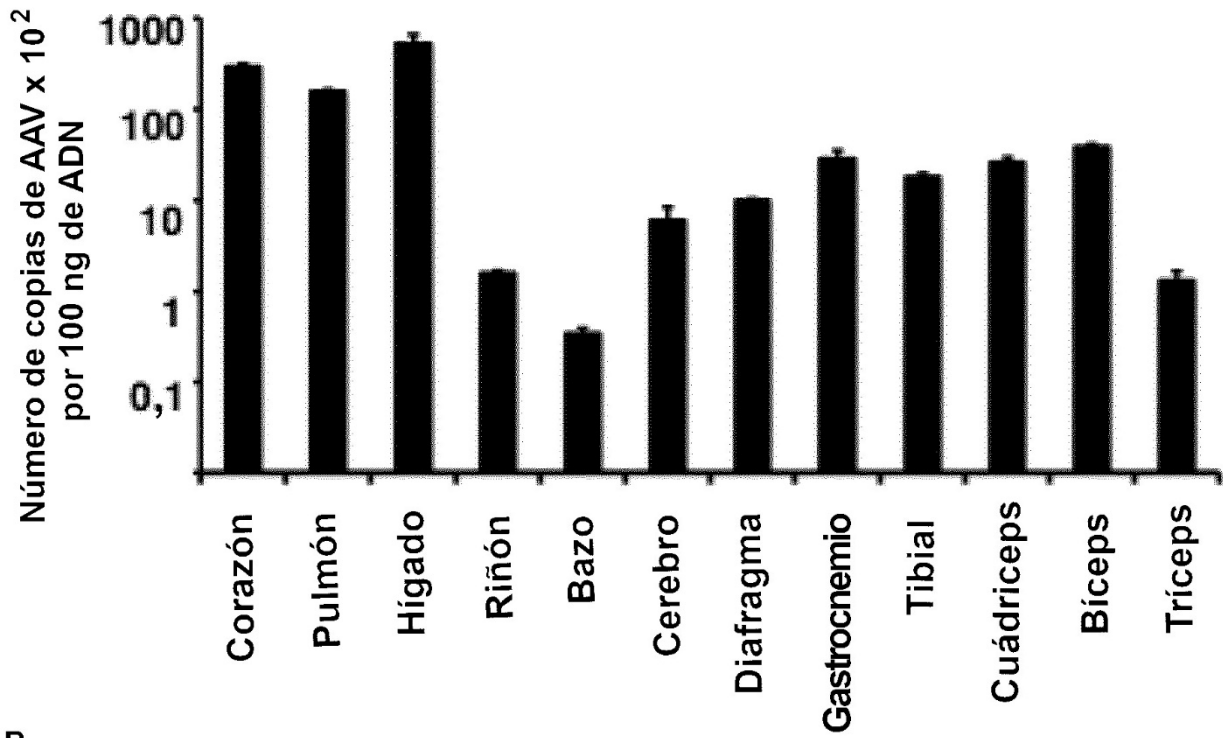


FIG. 13

A



B

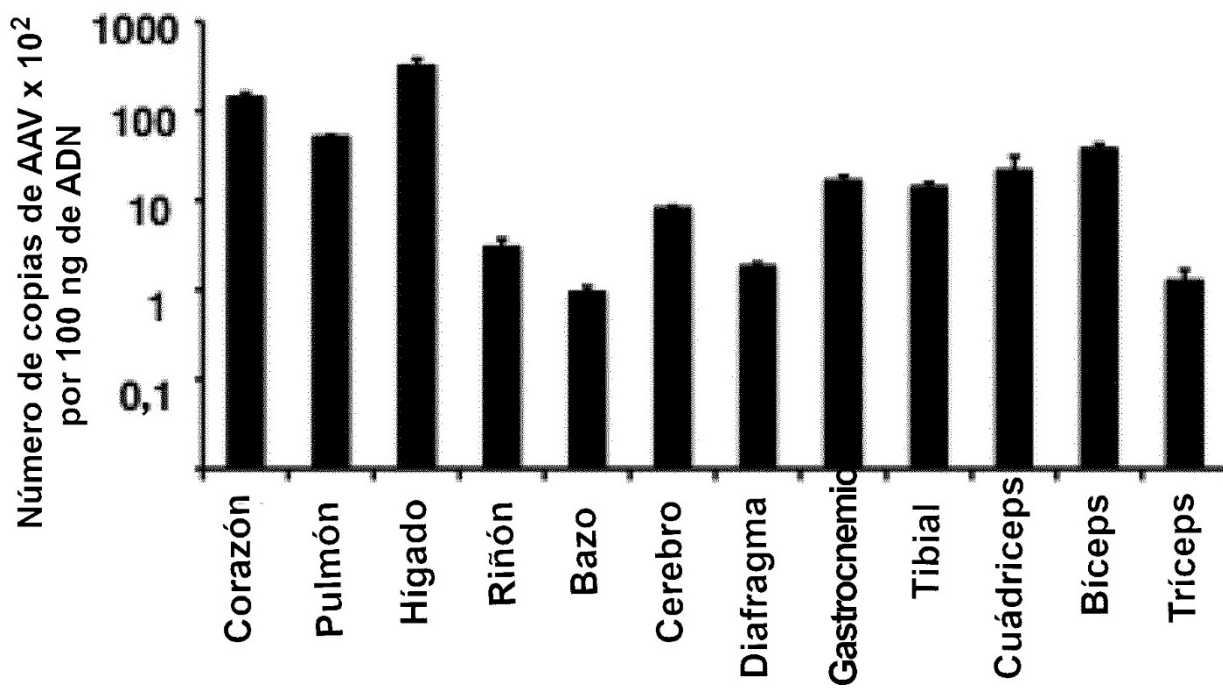


FIG. 14

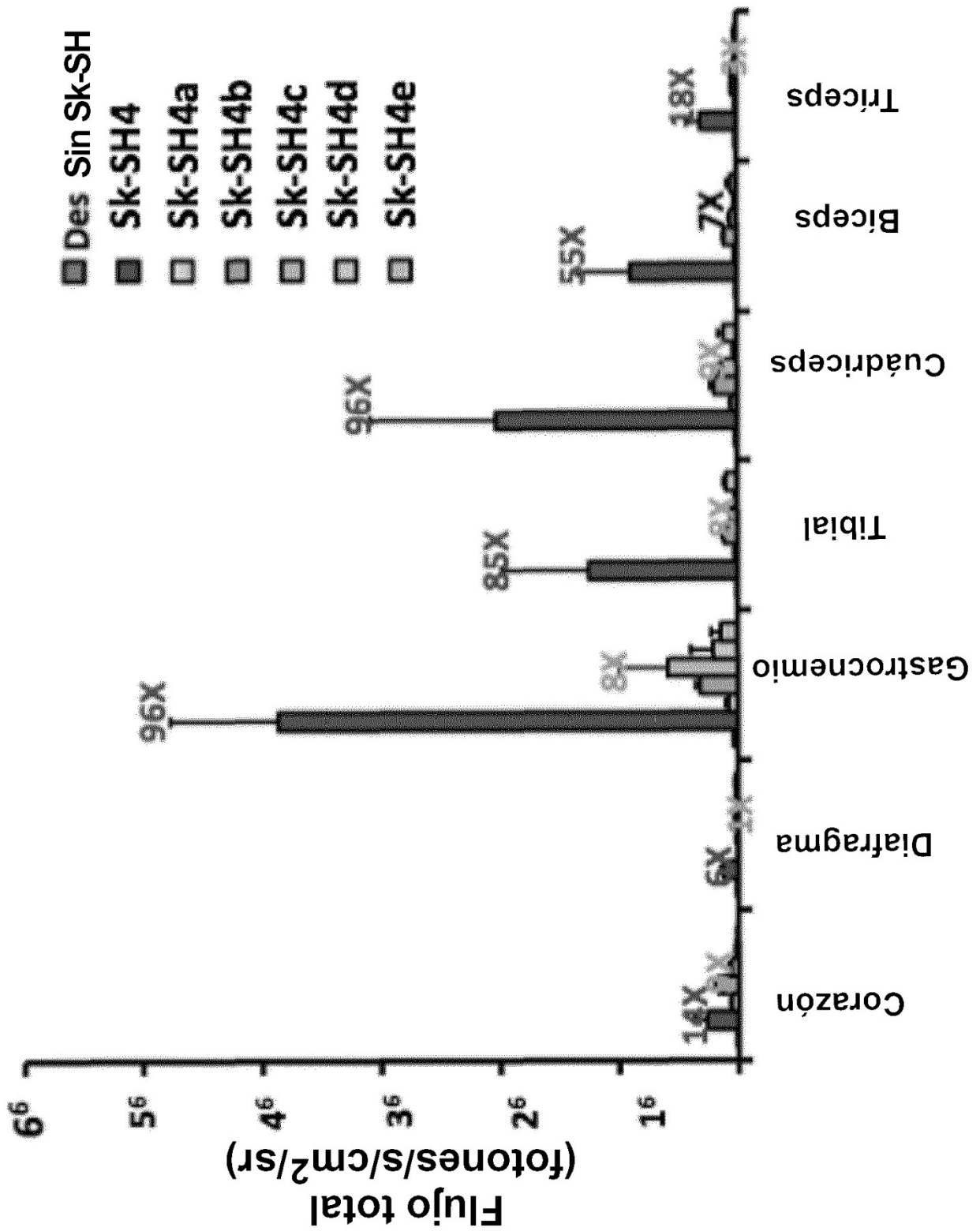
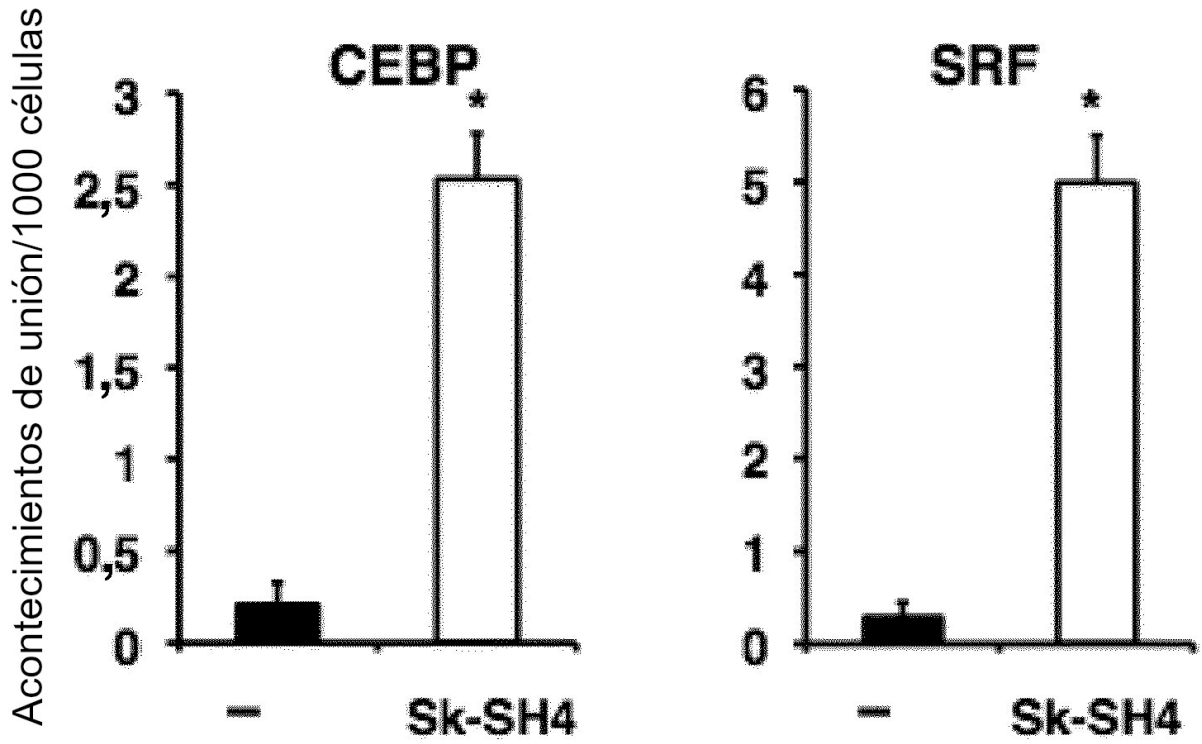


FIG. 15

A



B

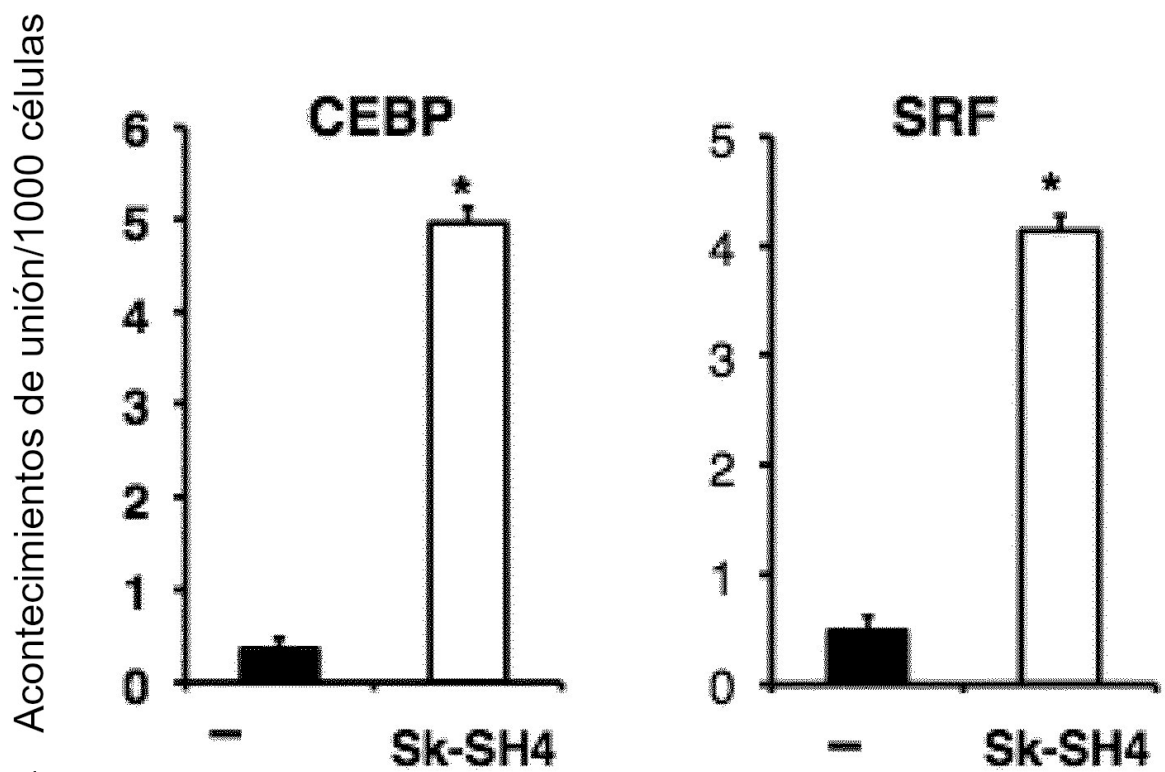
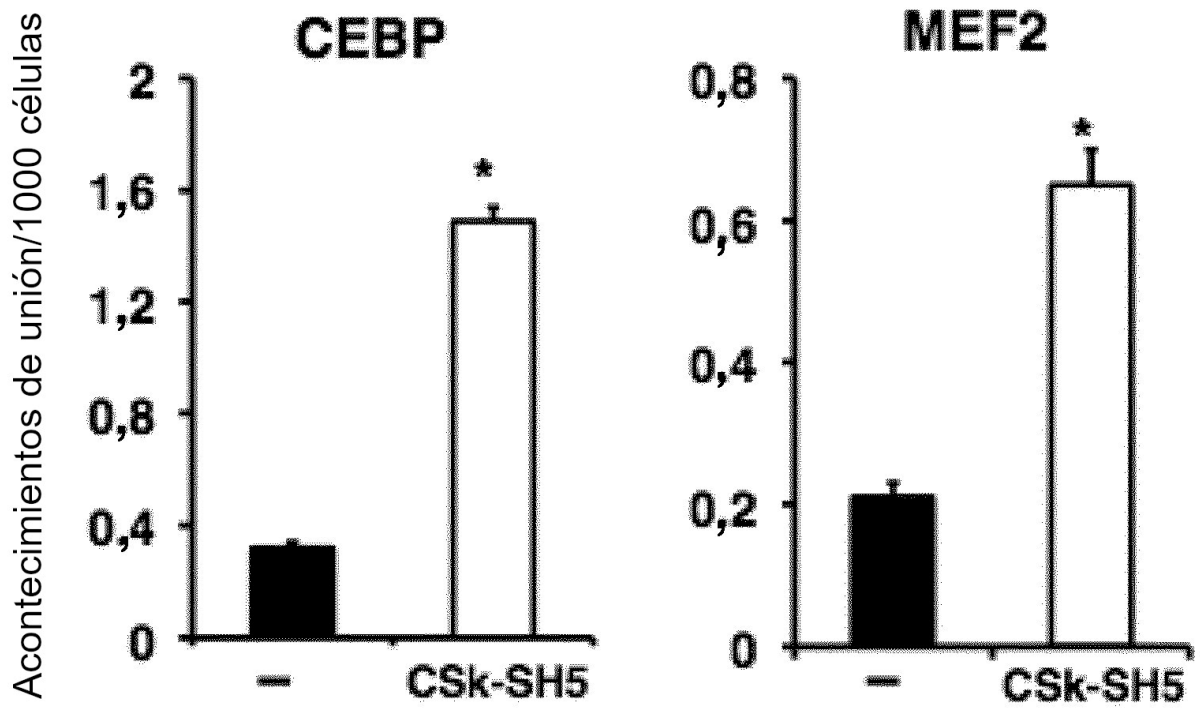


FIG. 15

C



D

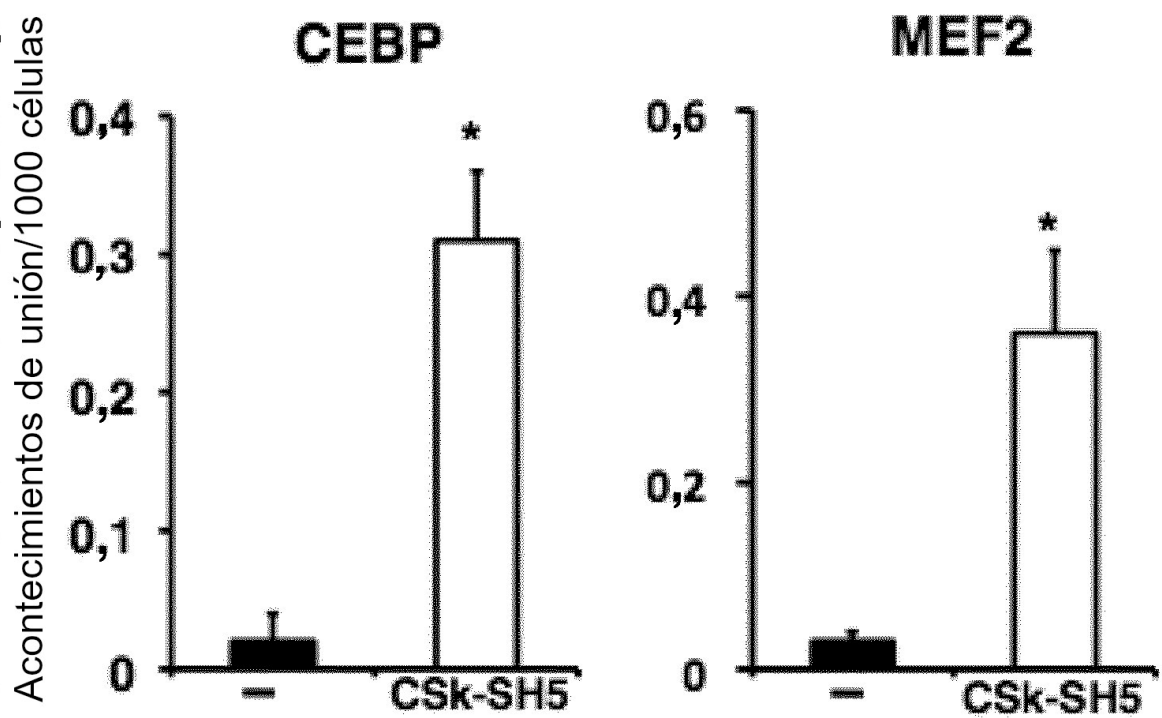


FIG. 16

A

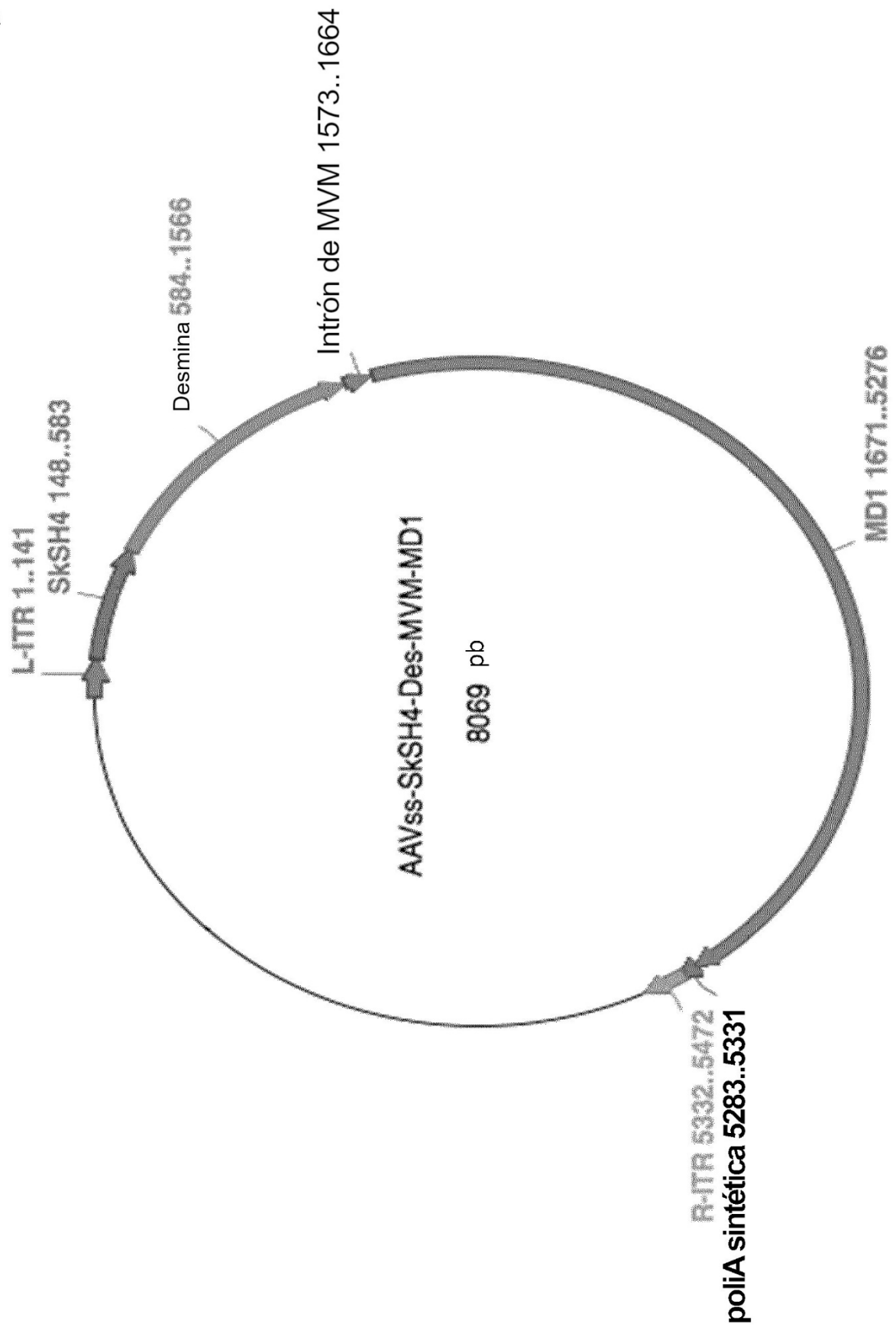


FIG. 16**B**

SEQ ID NO:44

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCAAAGCCCCGGGCGTCG
 GGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAAC
 TCCATCACTAGGGGTTCCTCGTACGTTCTGAGTCCCTCTAAGGTCCCTCACTCCCAACTCAG
 CCCCATGTCTGTCAATTCCCCTCAGTGTCTGATCTCCTTCTCCTCACCTTTCCCATCTC
 CCGTTTGACCCAAGCTTCCTGAGCTCTCCTCCCATTCCCCCTTTTGGAGTCCTCCTCCTCT
 CCCAGAACCCAGTAATAAGTGGGCTCCTCCCTGGCCTGGACCCCCGTGGTAACCTATAAG
 GCGAGGCAGCTGCTGTCTGAGGCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGCAGCTGCTA
 AGTTTAGGGTGGCTCCTTCTCTCTTCTTAGAGACAACAGGTGGCTGGGGCCTCAGTGCCCA
 GAAAAGAAAATGTCTTAGAGGTATCGGCATGGGCTGGAGGAGGGGGGACAGGGCAGGGGG
 AGGCATCTTCTCAGGACATCGGGTCCTAGAGGGACCTTGCTTCCTAGCTGGGCCCTTTCCT
 TCTCCTCTATAAATACCAGCTCTGGTATTTGCGCTTGGCAGCTGTTGCTGCTAGGGAGACG
 GCTGGCTTGACATGCATCTCCTGACAAAACACAAACCCGTGGTGTGAGTGGGTGTGGGCGG
 TGTGAGTAGGGGGATGAATCAGAGAGGGGGCGAGGGAGACAGGGGCGCAGGAGTCAGGCAA
 AGGCGATGCGGGGGTGCAGCTACACGCAGTTGGAAACAGTCGTCAGAAGATTCTGGAAACT
 ATCTTGCTGGCTATAAACTTGAGGGAAGCAGAAGGCCAACATTCCCTCCCAAGGGAAACTGA
 GGCTCAGAGTTAAAACCCAGGTATCAGTGATATGCATGTGCCCCGGCCAGGGTCACTCTCT
 GACTAACCGGTACCTACCCTACAGGCCTACCTAGAGACTCTTTTGAAAGGATGGTAGAGAC
 CTGTCCGGGCTTTGCCCACAGTCGTTGGAAACCTCAGCATTTTCTAGGCAACTTGTGCGAA
 TAAACACTTCGGGGGTCTTCTTGTTCATTCCAATAACCTAAAACCTCTCCTCGGAGAAA
 ATAGGGGGCCTCAAACAAACGAAATTCTCTAGCCCGCTTTCCCCAGGATAAGGCAGGCATC
 CAAATGGAAAAAAAGGGGCGCGGGGGGTCTCCTGTCAGCTCCTTGCCCTGTGAAACCCA
 GCAGGCCTGCCTGTCTTCTGTCTCTTGGGGCTGTCCAGGGGCGCAGGCCTCTTGCGGGGG
 AGCTGGCCTCCCCGCCCCCTCGCCTGTGGCCGCCCTTTTCTGTCAGGACAGAGGGATCCT
 GCAGCTGTCAGGGGAGGGGCGCCGGGGGGTGATGTCAGGAGGGCTACAAATAGTGCAGACA
 GCTAAGGGGCTCCGTCACCCATCTTCACATCCACTCCAGCCGGCTGCCCCGCCGCTGCCTC
 CTCTGTGCGTCCGCCCAGCCAGCCTCGTCCACGCCGCCACCTCTAGAAAGAGGTAAGGGTT
 TAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTGGAGCACCTGCCTGAAATC
 ACTTTTTTTCAGGTGGACGCGTGCCACCATGCTGTGGTGGGAGGAAGTGGAGGACTGCTA
 CGAGAGAGAGGACGTGCAGAAGAAAACCTTCACCAAGTGGGTGAACGCCCAGTTTCAAGCAAG
 TTCGGCAAGCAGCACATCGAGAACCTGTTTACGCGACCTGCAGGATGGCAGGAGACTGCTGG
 ATCTGCTGGAGGGACTGACCGGCCAGAAGCTGCCCCAAGGAGAAGGGCAGCACCAGAGTGCA
 CGCCCTGAACAACGTGAACAAGGCCCTGAGAGTGCTGCAGAACAAACAGTGGACCTGGTG
 AATATCGGCAGCACCGACATCGTGGACGGCAACCACAAGCTGACCCTGGGCCTGATCTGGA
 ACATCATCCTGCACTGGCAGGTGAAGAACGTGATGAAGAACATCATGGCCGGCCTGCAGCA
 GACCAACAGCGAGAAGATCCTGCTGAGCTGGGTGAGGCAGAGCACCAGAACTACCCCCAG
 GTGAACGTGATCAACTTCACCACCTCCTGGAGCGACGGCCTGGCCCTGAACGCCCTGATCC
 ACAGCCACAGACCCGACCTGTTTCACTGGAACAGCGTGGTGTGTCAGCAGAGCGCCACCCA
 GAGACTGGAGCACGCCTTCAACATCGCCAGATAACCAGCTGGGCATCGAGAAGCTGCTGGAC
 CCCGAGGACGTGGACACCACCTACCCCCGACAAGAAAAGCATCCTGATGTATATTACCTCTC
 TGTTTTAGGTGCTGCCCCAGCAGGTGTCCATCGAGGCCATCCAGGAAGTGGAAATGCTGCC
 CAGGCCCCCACCCTGTCCCTGGCCCAGGGCTATGAGAGAACCAGCAGCCCCAAGCCCAGA
 TTCAAGAGCACCGTGTCCCTGGCCCAGGGCTATGAGAGAACCAGCAGCCCCAAGCCCAGAT
 TCAAGAGCTACGCCTACACCCAGGCCGCCTACGTGACCACCTCCGACCCCCACCAGAAGCCC
 CTTCCCCAGCCAGCACCTGGAGGCCCCCCGAGGACAAGAGCTTCGGCAGCAGCCTGATGGAG

AGCGAAGTGAACCTGGACAGATACCAGACCGCCCTGGAGGAAGTGCTGTCTTGGCTGCTGT
 CCGCCGAGGACACCCCTGCAGGCCCCAGGGCGAGATCAGCAACGACGTGGAAGTGGTGAAGGA
 CCAGTTCCACACCCACGAGGGCTACATGATGGATCTGACCGCCCACCAGGGCAGAGTGGGC
 AATATCCTGCAGCTGGGCAGCAAGCTGATCGGCACCGGCAAGCTGAGCGAGGACGAGGAGA
 CCGAAGTGCAGGAGCAGATGAACCTGCTGAACAGCAGATGGGAGTGCCTGAGAGTGGCCAG
 CATGGAGAAGCAGAGCAACCTGCACCGCGTGCTGATGGACCTGCAGAACCAGAAGCTGAAG
 GAGCTGAACGACTGGCTGACCAAGACCGAGGAGCGGACCAGAAAGATGGAGGAGGAGCCCC
 TGGGCCCCGACCTGGAGGACCTGAAGAGACAGGTGCAGCAGCACAAAGTGCTGCAGGAGGA
 CCTGGAACAGGAGCAGGTGCGCGTGAACAGCCTGACCCACATGGTGGTCTGTTGGTGGACGAG
 AGCAGCGGCGACCACGCCACAGCCGCCCTGGAAGAGCAGCTGAAAGTGCTGGGCGACAGAT
 GGGCCAAACATCTGCCGGTGGACCGAGGACAGATGGGTGCTGCTGCAGGACATCCTGCTGAA
 GTGGCAGAGACTGACAGAGGAGCAGTGCCCTGTTTAGCGCCTGGCTGAGCGAGAAGGAGGAC
 GCCGTGAACAAGATCCACACCACCGGCTTCAAGGACCAGAACGAGATGCTGAGCAGCCTGC
 AGAAGCTGGCCGTGCTGAAGGCCGATCTGGAGAAGAAAAAGCAGAGCATGGGCAAGCTGTA
 CTCCCTGAAGCAGGACCTGCTGTCCACCCTGAAGAACAAGAGCGTGACCCAGAAAACCGAG
 GCCTGGCTGGACAATTTTCGCCCCGGTGTGGGACAATCTGGTGCAGAACTGGAGAAGAGCA
 CCGCCCAGATCAGCCAGGCCGTGACCACCACCCAGCCCAGCCTGACACAGACCACCGTGAT
 GGAGACCGTGACCACAGTGACCACCAGGGAGCAGATCCTGGTGAAGCACGCCACGAGGAG
 CTGCCCCCTCCCCCCCCCTCAGAAGAAGCGGCAGATCACAGTGGACACCCTGGAGAGACTGC
 AGGAGCTGCAGGAAGCCACCGACGAGCTGGACCTGAAGCTGAGACAGGCCGAAGTGATCAA
 GGGCAGCTGGCAGCCTGTGGGCGATCTGCTGATCGACAGCCTGCAGGACCACCTGGAGAAA
 GTGAAGGCCCTGCGGGGCGAGATCGCCCCCTGAAGGAGAATGTGAGCCACGTGAACGACC
 TGGCCAGACAGCTGACCACCCTGGGCATCCAGCTGAGCCCCTACAATCTGAGCACCCTGGA
 AGATCTGAACACCCGGTGGAAACTGCTGCAGGTGGCCGTGGAGGATAGAGTGAGGCAGCTG
 CACGAGGCCACAGAGACTTCGGCCCTGCCTCCCAGCACTTCCTGAGCACCAGCGTGCAGG
 GCCCCCTGGGAGAGAGCCATCTCCCCAACAAAGTGCCCTACTACATCAACCACGAGACCCA
 GACCACCTGCTGGGACCACCCTAAGATGACCGAGCTGTACCAGAGCCTGGCCGACCTGAAC
 AATGTGCGGTTTCAGCGCCTACAGAACCGCCATGAAGCTGCGGAGACTGCAGAAGGCCCTGT
 GCCTGGACCTGCTGAGCCTGAGCGCCGCGCTGCGACGCCCTGGACCAGCACAACTGAAGCA
 GAACGACCAGCCCATGGACATTCTGCAGATCATCAACTGCCTGACCACCATCTACGATCGG
 CTGGAGCAGGAGCACAAACCTGGTGAACGTGCCCTGTGCGTGGACATGTGCCTGAATT
 GGCTGCTGAACGTGTACGACACCGGCAGGACCGGCAGAATCAGAGTGCTGTCTTCAAGAC
 CGGCATCATCAGCCTGTGCAAGGCCACCTGGAGGATAAGTACCGCTACCTGTTCAAGCAG
 GTGGCCAGCAGCACCGGCTTCTGCGATCAGAGGAGACTGGGCCCTGCTGCTGCACGATAGCA
 TCCAGATCCCTAGGCAGCTGGGCGAAGTGGCCAGCTTTGGCGGCAGCAACATCGAGCCCTC
 TGTGAGGAGCTGCTTCCAGTTCGCCAACAAAGCCCCGAGATCGAGGCCGCCCTGTTCTCTG
 GATTGGATGAGGCTGGAGCCCCAGAGCATGGTGTGGCTGCCTGTGCTGCACAGAGTGGCCG
 CCGCCGAGACCGCCAAGCACAGGCCAAGTGCAACATCTGCAAGGAGTGCCCCATCATCGG
 CTTCCGGTACAGGAGCCTGAAGCACTTCAACTACGACATCTGCCAGAGCTGCTTTTTTCAGC
 GGCAGAGTGGCCAAGGGCCACAAGATGCACTACCCCATGGTGGAGTACTGCACCCCCACCA
 CCTCCGGCGAGGATGTGAGAGACTTCGCCAAAGTGCTGAAGAATAAGTTCCGGACCAAGCG
 GTACTTTGCCAAGCACCCCAGGATGGGCTACCTGCCCCTGCAGACCGTGCTGGAGGGCGAC
 AACATGGAGACCGACACCATGTGATGATGACTCGAGAATAAAAGATCTTTATTTTCATTAG
 ATCTGTGTGTGGTTTTTTGTGTGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCT
 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCGGGCTTTGCC
 CGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGGGGCGCCTGATGCGGTATT
 TTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATACGTCAAAGCAACCATAGTACGCG
 CCCTGTAGCGGCGCATTAAGCGCGGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACAC
 TTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCTCTCGCCACGTTTCG

CGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTA
CGGCACCTCGACCCCAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCT
GATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGAATCTTGTT
CCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTG
CCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTA
ACAAAATATTAACGTTTACAATTTTATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGC
ATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGCTCTG
CTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCTAGAGGT
TTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATA
GGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTTCGGGGAAATGTG
CGCGGAACCCCTATTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGAC
AATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTT
CCGTGTCGCCCTTATTCCTTTTTTTCGGGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAA
ACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAAC
TGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTTCCAATGAT
GAGCACTTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAG
CAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAG
AAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAG
TGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCT
TTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATG
AAGCCATAACCAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCG
CAAATATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATG
GAGGCGGATAAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTG
CTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGA
TGGTAAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAA
CGAAATAGACAGATCGCTGAGATAGGTGCCCTCACTGATTAAGCATTTGGTAACTGTCAGACC
AAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTTAATTTAAAAGGATCTA
GGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCAC
TGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCG
TAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCA
AGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACT
GTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACAT
ACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTAC
CGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGGCTGAACGGGGGGT
TCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTG
AGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGG
CAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTAT
AGTCCTGTGCGGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGG
GGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTTACGGTTCCTGGCCTTTTGCTG
GCCTTTTGCTCACATGT

FIG. 16

C

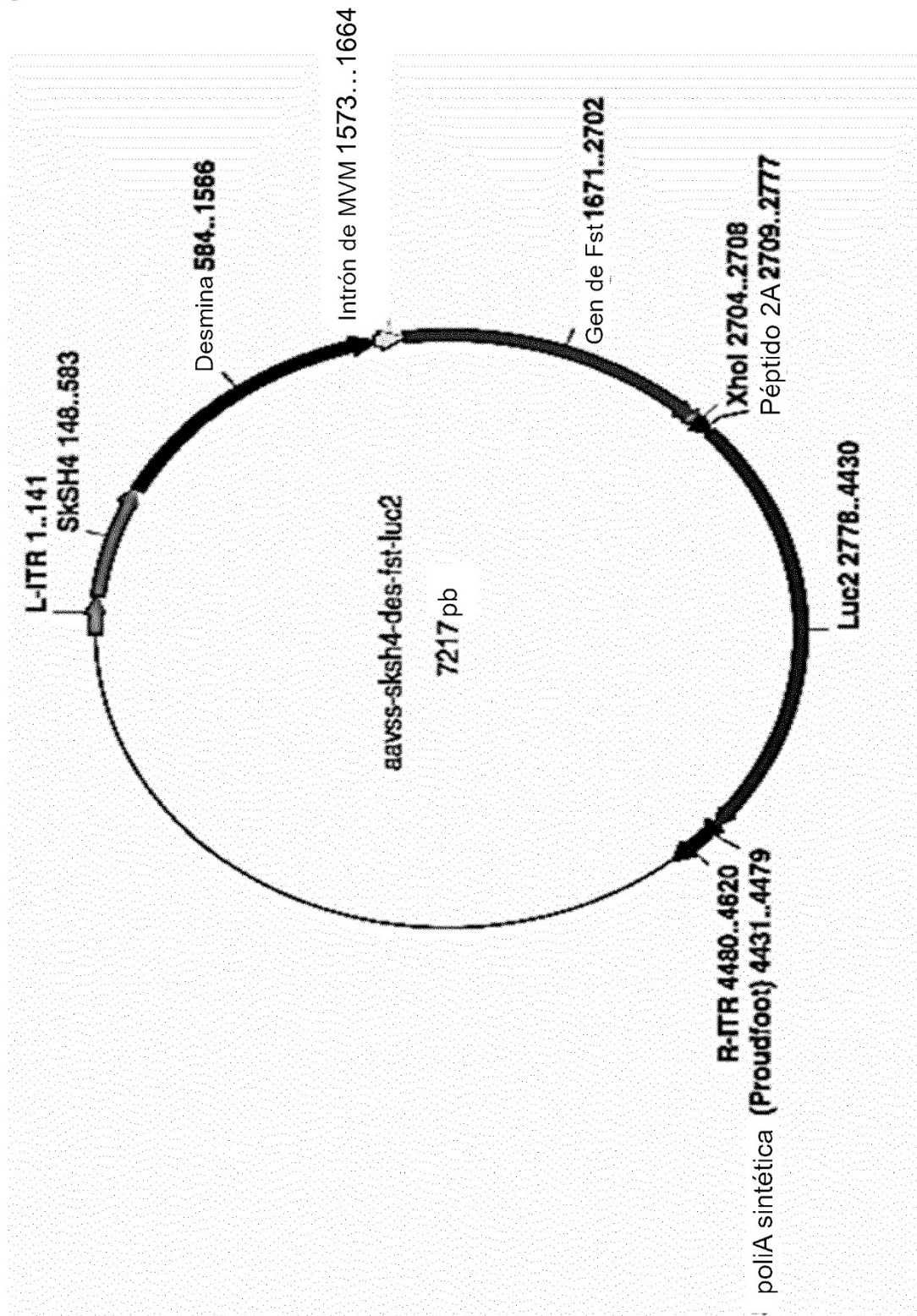


FIG. 16**D**

SEQ ID NO:45

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCAAAGCCCCGGGCGTCG
 GGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAAC
 TCCATCACTAGGGGTTCCTCGTACGTTCTGAGTCCCTCTAAGGTCCCTCACTCCCAACTCAG
 CCCCATGTCTGTCAATTCCCCTCAGTGTCTGATCTCCTTCTCCTCACCTTTCCCATCTC
 CCGTTTGACCCAAGCTTCCTGAGCTCTCCTCCCATTCCCCCTTTTGGAGTCCTCCTCCTCT
 CCCAGAACCCAGTAATAAGTGGGCTCCTCCCTGGCCTGGACCCCCGTGGTAACCCATATAAG
 GCGAGGCAGCTGCTGTCTGAGGCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGCAGCTGCTA
 AGTTTAGGGTGGCTCCTTCTCTCTTCTTAGAGACAACAGGTGGCTGGGGCCTCAGTGCCCA
 GAAAAGAAAATGTCTTAGAGGTATCGGCATGGGCCTGGAGGAGGGGGGACAGGGCAGGGGG
 AGGCATCTTCTCAGGACATCGGGTCTTAGAGGGACCTTGCTTCCTAGCTGGGCCTTTCTCT
 TCTCCTCTATAAATACCAGCTCTGGTATTTTCGCCTTGGCAGCTGTTGCTGCTAGGGAGACG
 GCTGGCTTGACATGCATCTCCTGACAAAACACAAACCCGTGGTGTGAGTGGGTGTGGGCGG
 TGTGAGTAGGGGGATGAATCAGAGAGGGGGCGAGGGAGACAGGGGGCGCAGGAGTCAGGCAA
 AGGCGATGCGGGGGTGCAGCTACACGCAGTTGGAAACAGTCGTCAGAAGATTCTGGAAACT
 ATCTTGCTGGCTATAAACTTGAGGGAAGCAGAAGGCCAACATTCCCTCCCAAGGGAAACTGA
 GGCTCAGAGTTAAAACCCAGGTATCAGTGATATGCATGTGCCCCGGCCAGGGTCACTCTCT
 GACTAACCGGTACCTACCCTACAGGCCTACCTAGAGACTCTTTTGAAAGGATGGTAGAGAC
 CTGTCCGGGCTTTGCCCCACAGTCGTTGGAAACCTCAGCATTTTCTAGGCAACTTGTGCGAA
 TAAAACACTTCGGGGGTCTTCTTGTTTCATTCCAATAACCTAAAACCTCTCCTCGGAGAAA
 ATAGGGGGCCTCAAACAAACGAAATTCTCTAGCCCGCTTTCCCCAGGATAAGGCAGGCATC
 CAAATGGAAAAAAGGGGGCCGGCCGGGGGTCTCCTGTCAGCTCCTTGCCCTGTGAAACCCA
 GCAGGCCTGCCTGTCTTCTGTCTCTTGGGGCTGTCCAGGGGGCGCAGGCCTCTTGCGGGGG
 AGCTGGCCTCCCCGCCCCCTCGCCTGTGGCCGCCCTTTTCTGTCAGGACAGAGGGATCCT
 GCAGCTGTCAGGGGAGGGGCGCCGGGGGGTGATGTCAGGAGGGCTACAAATAGTGCAGACA
 GCTAAGGGGCTCCGTCACCCATCTTCACATCCACTCCAGCCGGCTGCCCCGCCGCTGCCTC
 CTCTGTGCGTCCGCCCAGCCAGCCTCGTCCACGCCGCCACCTCTAGAAAGAGGTAAGGGTT
 TAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTGGAGCACCTGCCTGAAATC
 ACTTTTTTTTCAGGTTGGACGCGTATGGTCCGCGCGAGGCACCAGCCGGGTGGGCTTTGCCT
 CCTGCTGCTGCTGCTCTGCCAGTTCATGGAGGACCGCAGTGCCAGGCTGGGAACCTGCTGG
 CTCCGTCAAGCGAAGAACGGCCGCTGCCAGGTCTGTACAAGACCGAACTGAGCAAGGAGG
 AGTGCTGCAGCACCGGCCGGCTGAGCACCTCGTGGACCGAGGAGGACGTGAATGACAACAC
 ACTCTTCAAGTGGATGATTTTCAACGGGGGCGCCCCCAACTGCATCCCCCTGTAAAGAAACG
 TGTGAGAACGTGGACTGTGGACCTGGGAAAAAATGCCGAATGAACAAGAAGACAAACCCC
 GCTGCGTCTGCGCCCCGGATTGTTCCAACATCACCTGGAAGGGTCCAGTCTGCGGGCTGGA
 TGGGAAAACCTACCGCAATGAATGTGCACTCCTAAAGGCAAGATGTAAAGAGCAGCCAGAA
 CTGGAAGTCCAGTACCAAGGCAGATGTAAAAAGACTTGTTCGGGATGTTTTCTGTCCAGGCA
 GCTCCACATGTGTGGTGGACCAGACCAATAATGCCTACTGTGTGACCTGTAATCGGATTTG
 CCCAGAGCCTGCTTCTCTGAGCAATATCTCTGTGGGAATGATGGAGTCACCTACTCCAGT
 GCCTGCCACCTGAGAAAGGCTACCTGCCTGCTGGGCAGATCTATTGGATTAGCCTATGAGG
 GAAAGTGTATCAAAGCAAAGTCTGTGAAGATATCCAGTGCAGTGGTGGGAAAAAATGTTT
 ATGGGATTTCAAGGTTGGGAGAGGCCGGTGTTCCCTCTGTGATGAGCTGTGCCCTGACAGT
 AAGTCGGATGAGCCTGTCTGTGCCAGTGACAATGCCACTTATGCCAGCGAGTGTGCCATGA
 AGGAAGCTGCCTGCTCCTCAGGTGTGCTACTGGAAGTAAAGCACTCCGGATCTTGCAACTC
 CATTTTCGGAAGACACCGAGGAAGAGGAGGAAGATGAAGACCAGGACTACAGCTTTCCTATA

TCTTCTATTCTAGAGTGGGTTCGAGAGGTCCGGCGGCGGAGAGGGCAGAGGAAGTCTTCTAA
 CATGCGGTGACGTGGAGGAGAATCCCGGCCCAATGGAAGATGCCAAAAACATTAAGAAGGG
 CCCAGCGCCATTCTACCCACTCGAAGACGGGACCGCCGGCGAGCAGCTGCACAAAGCCATG
 AAGCGCTACGCCCTGGTGCCCGGCACCATCGCCTTTACCGACGCACATATCGAGGTGGACA
 TTACCTACGCCGAGTACTTCGAGATGAGCGTTTCGGCTGGCAGAAGCTATGAAGCGCTATGG
 GCTGAATACAAACCATCGGATCGTGGTGTGCAGCGAGAATAGCTTGCAGTTCTTCATGCC
 GTGTTGGGTGCCCTGTTTCATCGGTGTGGCTGTGGCCCCAGCTAACGACATCTACAACGAGC
 GCGAGCTGCTGAACAGCATGGGCATCAGCCAGCCCACCGTCGTATTCGTGAGCAAGAAAGG
 GCTGCAAAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAAAAGATCATCATCATG
 GATAGCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTCGTGACTTCCCATTTGC
 CACCCGGCTTCAACGAGTACGACTTCGTGCCCCGAGAGCTTCGACCGGGACAAAACCATCGC
 CCTGATCATGAACAGTAGTGGCAGTACCGGATTGCCCAAGGGCGTAGCCCTACCGCACCGC
 ACCGCTTGTGTCCGATTTCAGTCATGCCCCGCGACCCCATCTTCGGCAACCAGATCATCCCCG
 ACACCGCTATCCTCAGCGTGGTGCCATTTACCACGGCTTCGGCATGTTACCACGCTGGG
 CTACTTGATCTGCGGCTTTTCGGGTCTGTGCTCATGTACCGCTTCGAGGAGGAGCTATTCTTG
 CGCAGCTTGCAAGACTATAAGATTCAATCTGCCCTGCTGGTGCCCACACTATTTAGCTTCT
 TCGCTAAGAGCACTCTCATCGACAAGTACGACCTAAGCAACTTGCACGAGATCGCCAGCGG
 CGGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGCCGTGGCCAAACGCTTCCACCTACCAGGC
 ATCCGCCAGGGCTACGGCCTGACAGAAACAACCAGCGCCATTCTGATCACCCCGAAGGGG
 ACGACAAGCCTGGCGCAGTAGGCAAGGTGGTGCCCTTCTTCGAGGCTAAGGTGGTGGACTT
 GGACACCGGTAAGACACTGGGTGTGAACCAGCGCGGCGAGCTGTGCGTCCGTGGCCCCATG
 ATCATGAGCGGCTACGTTAACAACCCCGAGGCTACAAACGCTCTCATCGACAAGGACGGCT
 GGCTGCACAGCGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGTGGACCG
 GCTGAAGAGCCTGATCAAATACAAGGGCTACCAGGTAGCCCCAGCCGAACCTGGAGAGCATC
 CTGCTGCAACACCCCAACATCTTCGACGCGCGGGGTGCGCCGGCCTGCCCGACGACGATGCCG
 GCGAGCTGCCCCCGCAGTCGTCTGTGCTGGAACACGGTAAAACCATGACCGAGAAGGAGAT
 CGTGGACTATGTGGCCAGCCAGGTTACAACCGCCAAGAAGCTGCGCGGTGGTGTGTGTTT
 GTGGACGAGGTGCCTAAAGGACTGACCGGCAAGTTGGACGCCCCGAAGATCCGCGAGATT
 TCATTAAGGCCAAGAAGGGCGGCAAGATCGCCGTGTAAATAAAAAGATCTTTATTTTTCATT
 AGATCTGTGTGTTGGTTTTTTTGTGTGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCT
 CTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTG
 CCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGGGGCGCCTGATGCGGTA
 TTTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATACGTCAAAGCAACCATAGTACG
 CGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTAC
 ACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTCTTCCCTTCTTCTCGCCACGTTT
 GCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTT
 TACGGCACCTCGACCCCCAAAAAAGTTGATTTGGGTGATGGTTACGTTAGTGAGTGGGCCATCGCC
 CTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTG
 TTCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTT
 TGCCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTT
 TAACAAAAATATTAACGTTTACAATTTTATGGTGCACCTCTCAGTACAATCTGCTCTGATGCC
 GCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTC
 TGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAG
 GTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTA
 TAGGTAAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATG
 TGCGCGGAACCCCTATTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAG
 ACAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACAT
 TTCCGTGTGCCCCCTATTCCCTTTTTTGCGGCATTTTGCCCTTCTGTTTTTGCTCACCCAG
 AAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGA

ACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATG
 ATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAG
 AGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCAC
 AGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATG
 AGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCG
 CTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAA
 TGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTG
 CGCAAACCTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGA
 TGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTAT
 TGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTTGCAGCACTGGGGCCA
 GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATG
 AACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTTGGTAACTGTCAGA
 CCAAGTTTACTCATATATACTTTTAGATTGATTTTAAAACCTCATTTTTTAATTTAAAAGGATC
 TAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAAATCCCTTAACGTGAGTTTTTCGTTCC
 ACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCG
 CGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGAT
 CAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATA
 CTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTAC
 ATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTT
 ACCGGGTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCTGGGCTGAACGGGGG
 GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCG
 TGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGC
 GGCAGGGTCTGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTT
 ATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGG
 GGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGC
 TGGCCTTTTGCTCACATGT

FIG. 17

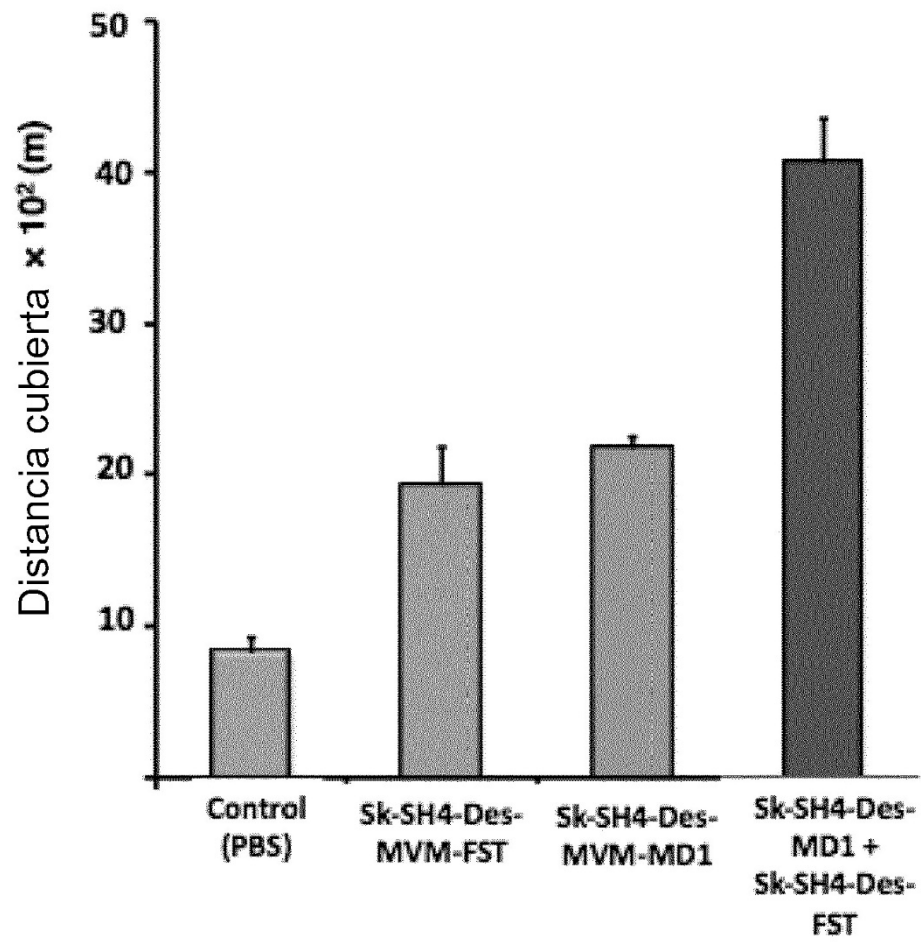
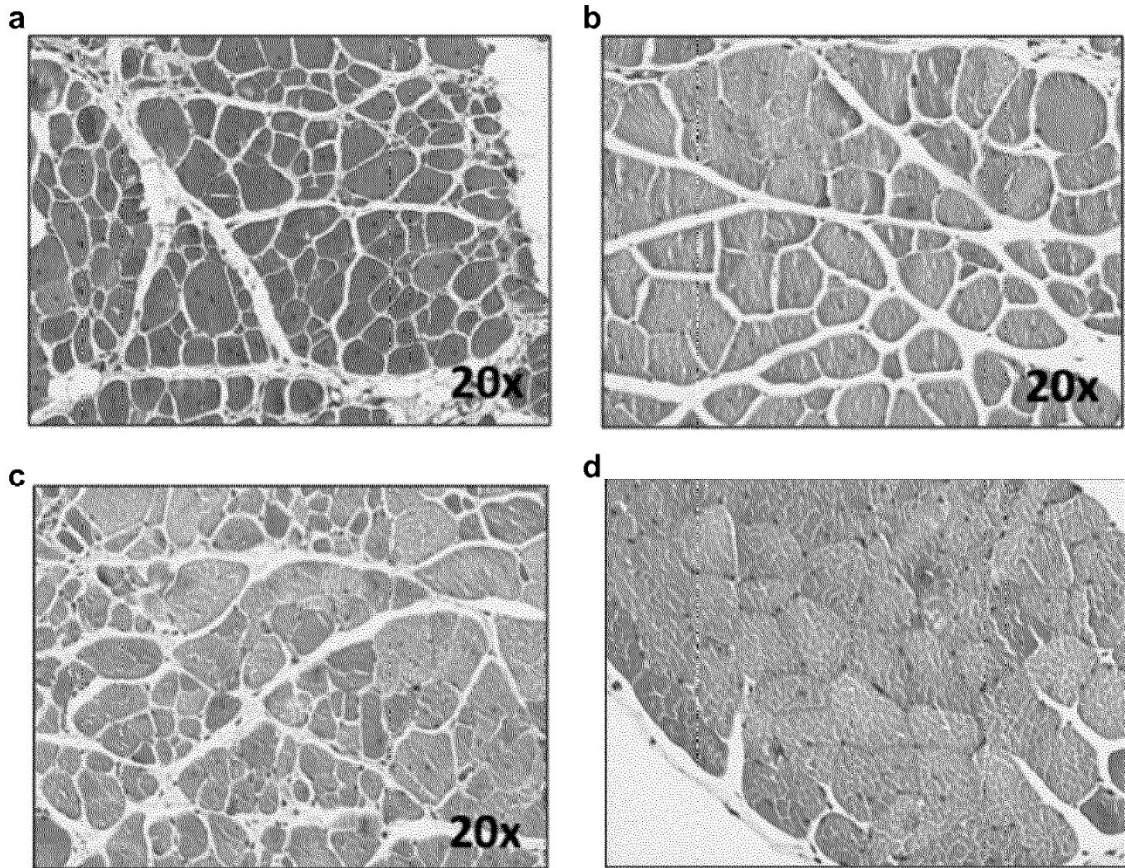


FIG. 18

A



B

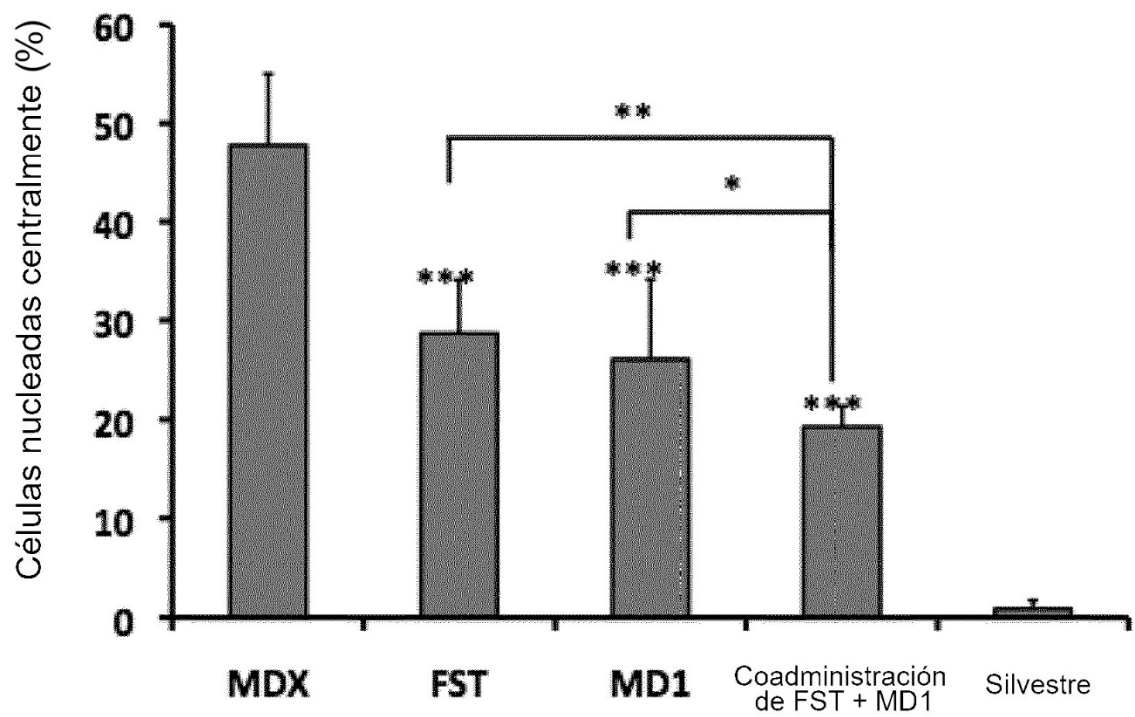
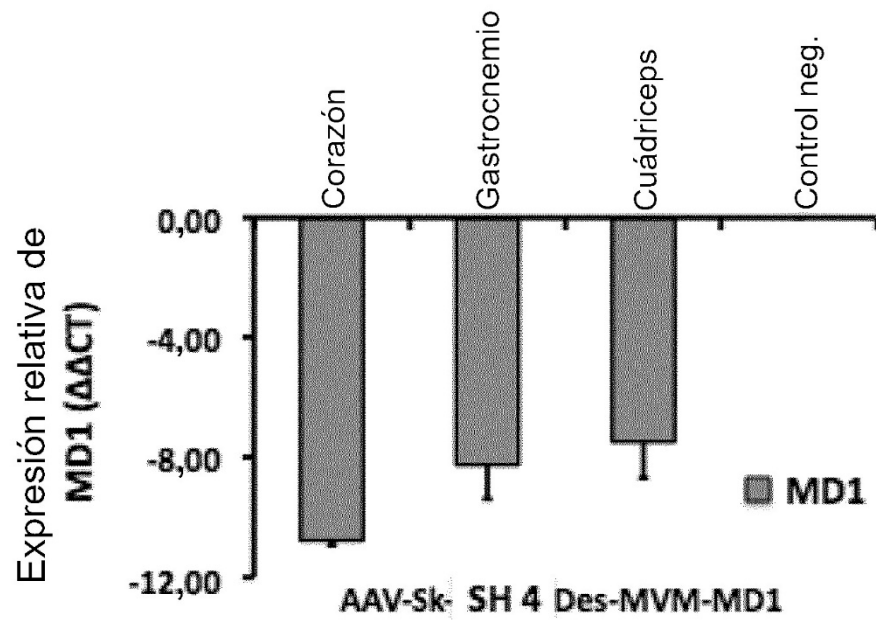


FIG. 19

A



B

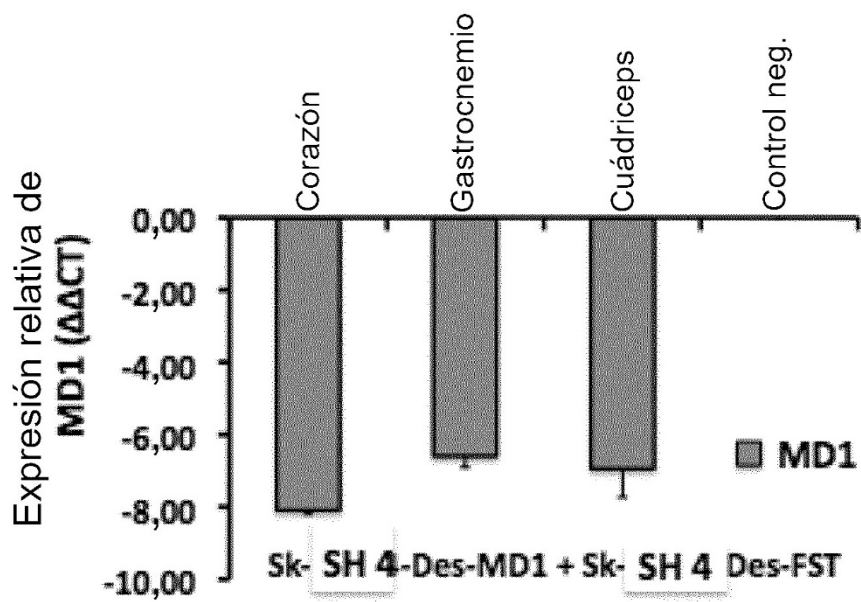


FIG. 19

C

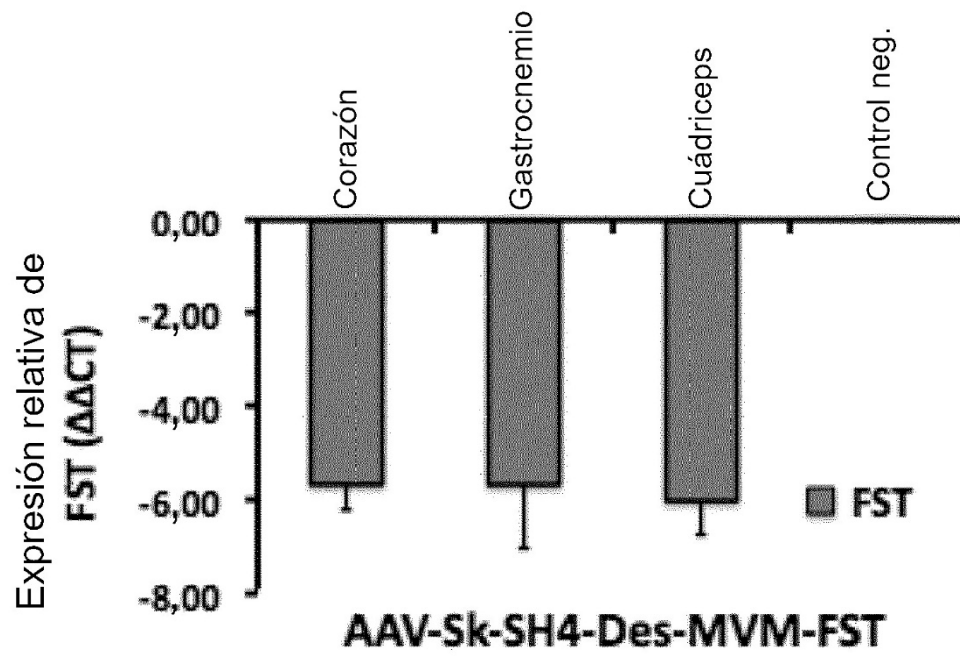


FIG. 20

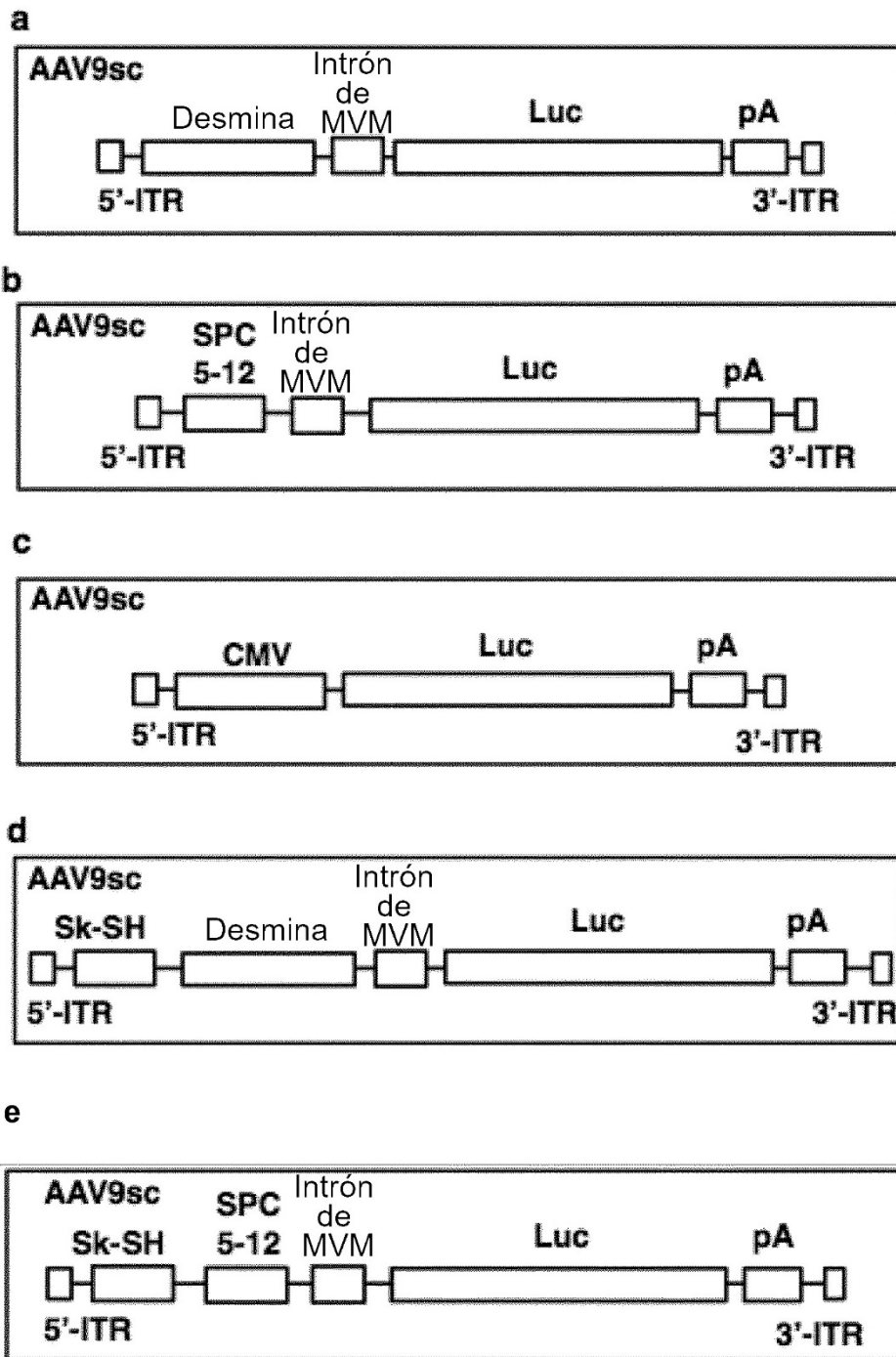


FIG. 21

