

TESIS DOCTORAL

EDICIÓN GÉNICA PARA TERAPIA GÉNICA DEL
SÍNDROME DE WISKOTT-ALDRICH

Sabina Sánchez Hernández

Directores de la tesis:

Dr. Karim Benabdel Lah El Khlanji

Dr. Francisco Martín Molina

Granada, 2019

Editor: Universidad de Granada. Tesis Doctorales
Autor: Sabina Sánchez Hernández
ISBN: 978-84-1306-297-6
URI: <http://hdl.handle.net/10481/56858>

El doctorando / The *doctoral candidate* [**Sabina Sánchez Hernández**] y los directores de la tesis / and the thesis supervisor/s: [**Francisco Martín y Karim Benabdellah**]

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

/

Guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisor/s and, as far as our knowledge reaches, in the performance of the work, the rights of other authors to be cited (when their results or publications have been used) have been respected.

Lugar y fecha / Place and date:

Granada, 10 Julio, 2019

Director/es de la Tesis / *Thesis supervisor/s*;

Doctorando / *Doctoral candidate*:

Firma / Signed

Firma / Signed

ÍNDICE

ÍNDICE	3
ABREVIATURAS	7
RESUMEN	14
INTRODUCCIÓN	21
1.1 TERAPIA GÉNICA	23
1.2 EDICIÓN GENÓMICA	23
1.2.1 Plataformas de EG independientes de nucleasa	26
1.2.1.1 Oligonucleótidos formadores de triplex.....	26
1.2.1.2 Vectores recombinantes adeno-asociados (rAAV)	27
1.2.2 Plataformas de EG con actividad endonucleasa.....	27
1.2.2.1 Meganucleasas.....	28
1.2.2.2 ZFNs	30
1.2.2.3 TALENs.....	33
1.2.2.4 CRISPR/Cas.....	34
1.3 MECANISMOS DE REPARACIÓN DEL DNA	36
1.3.1 NHEJ.....	36
1.3.2 Recombinación Homóloga	39
1.4 LIMITACIONES DE LA EG: EFECTOS OFF TARGET	40
1.4.1 Métodos sesgados de detección de sitios off target	41
1.4.2 Métodos no sesgados de detección de off target	43
1.5 EDICIÓN GENÓMICA EN TERAPIA GÉNICA	46
1.5.1 Vectores para terapia génica	47
1.5.1.1 Vectores Adenovirales (AdV)	48
1.5.1.2 Vectores Adeno-Asociados (rAAV)	49
1.5.1.3 Vectores retrovirales (RVs)	51
1.5.1.3.1 Vectores lentivirales (LVs).....	52
1.5.1.3.2 Vectores lentivirales no integrativos (IDLVs)	53
1.6 SÍNDROME DE WISKOTT-ALDRICH (WAS)	55
1.6.1 Terapias para el tratamiento de WAS	57

JUSTIFICACIÓN	61
HIPÓTESIS	66
OBJETIVOS	68
MATERIAL Y MÉTODOS	72
1.7 TIPOS CELULARES Y CONDICIONES DE CULTIVO	74
1.7.1 Línea celular HEK 293T.....	74
1.7.2 Línea celular K562.....	74
1.7.3 Línea celular Jurkat.....	74
1.7.4 Células progenitoras neuronales humanas (NPCs).....	74
1.7.4.1 Diferenciación neuronal in vitro y cultivo de neuronas diferenciadas	75
1.7.5 Células troncales pluripotentes inducidas (iPSCs).....	75
1.7.6 Cultivos celulares primarios de mucosa oral humana y fibroblastos de piel.....	75
1.7.7 Aislamiento de células troncales hematopoyéticas CD34+ (HSCs).....	76
1.7.7.1 Cultivo de HSCs (CD34+).....	76
1.7.8 Linfocitos T primarios aislados a partir de sangre periférica movilizada.....	77
1.8 SISTEMAS UTILIZADOS PARA ENTREGA DE NE Y DNA DONADORES.....	77
1.8.1 Producción y titulación de LV e IDLV de segunda generación.....	78
1.8.1.1 Transducción de diferentes tipos celulares con LV o IDLV.....	79
1.8.2 Formación de ribonucleoproteína Cas9/gRNA in vitro	80
1.8.3 Nucleofección	80
1.9 GENERACIÓN DE MODELOS CELULARES PARA ESTUDIAR EFICACIA Y SEGURIDAD DE LA EG	81
1.9.1 Generación de la línea celular reportera K562 SE- δ GF-I ₁ -P	81
1.9.2 Generación de la línea celular reportera K562 SEWAS84	82
1.10 GENERACIÓN DE DONADORES PARA RH.....	84
1.10.1 Donadores para el modelo celular K562 SE- δ GF-I ₁ -P.....	84
1.10.1.1 SEED.....	84
1.10.1.2 SECdR, SECdR_IS2, y SECdR_WI ₁ R.....	84
1.10.2 Donador para el modelo celular K562 SEWAS84.....	85
1.10.2.1 SdR*GFP	85
1.10.3 Donador para el locus <i>WAS</i> endógeno	85
1.11 CARACTERIZACIÓN DE EFICACIA DE LOS IDLVs	85

1.11.1	Perfil de formas extracromosómicas.....	85
1.11.2	RT-qPCR para determinación de niveles de mRNA.....	86
1.11.3	FISH.....	86
1.12	CARACTERIZACIÓN DE LA EFICACIA Y ESPECIFICIDAD DE CORTE DE LAS NE.....	87
1.12.1	Cuantificación de eficiencia de corte en sitio diana	87
1.12.2	Análisis de efectos off target.....	88
1.12.3	Tinción intracelular de WASP.....	92
1.12.4	Análisis estadísticos	93
RESULTADOS.....		96
1.13	OBJETIVO 1: MEJORA DE LA EFICACIA DE EDICIÓN GENÓMICA EN CÉLULAS T Y HSCs.....	97
1.13.1	Evaluación de la eficiencia y especificidad de corte en el locus <i>WAS</i> utilizando diferentes métodos de entrega del sistema CRISPR/Cas9 en células K562.....	97
1.13.2	Evaluación de la eficiencia y especificidad de corte en el locus <i>WAS</i> utilizando diferentes métodos de entrega del sistema CRISPR/Cas9 en células T y HSCs.....	102
1.13.3	Desarrollo de nuevos sistemas de entrega más eficaces basados en IDLVs.....	106
1.13.3.1	La inclusión del elemento IS2 en el LTR de IDLVs mejora sus niveles de expresión en las líneas celulares K562y 293T.....	109
1.13.3.2	El elemento IS2 no previene el silenciamiento de IDLVs debido a HDAC.....	115
1.13.3.3	La inserción del elemento IS2 en IDLVs no afecta el empaquetamiento de RNA en las partículas virales, pero reduce la cantidad de episomas en las células diana.....	117
1.13.3.4	Episomas de IDLVs que contienen el elemento IS2 producen mayores niveles de RNAm, y muestran distinta localización en el núcleo.....	119
1.13.3.5	La inserción del elemento IS2 en IDLVs reduce la formación de círculos 2-LTR, de baja expresión	123
1.13.3.6	La inclusión del elemento IS2 mantiene su efecto positivo en IDLV que expresan el transgén a través del promotor fisiológico de WASP.....	127
1.13.3.7	El efecto final de la inclusión de IS2 en IDLVs depende de la célula diana..	129
1.13.3.7.1	La inclusión de IS2 en IDLV mejora los niveles de expresión en NPCs y neuronas diferenciadas	133

1.13.3.7.2	IS2-IDLVs mejoran la eficiencia de transducción y los niveles de expresión en iPSCs.....	135
1.13.3.8	Inclusión de IS2 en IDLVs para entrega de Cas9 en HSCs y K562.....	137
1.14	OBJETIVO 2: DESARROLLO DE MODELOS PARA ESTUDIAR EFICIENCIA Y SEGURIDAD DE LA REPARACIÓN GÉNICA MEDIANTE CRISPR/CAS9.....	139
1.14.1	Generación de la línea reportera K562 SE- δ GF-I ₁ -P.....	139
1.14.1.1	Análisis de la eficacia y seguridad de EG en el modelo K562- δ GF-I ₁ -P.....	141
1.14.1.1.1	Comparativa de la eficacia y especificidad de la EG de diferentes sistemas de entrega del sistema CRISPR/Cas9 utilizando el donador SEED.....	144
1.14.1.1.2	Comparativa de la eficacia y especificidad de la EG de diferentes donadores y diferentes sistemas de entrega.....	148
1.14.1.2	Generación de la línea reportera K562 SEWAS84.....	153
1.14.1.2.1	Diseño del donador SdR*GFP.....	155
1.14.1.3	Comparativa de la eficacia y especificidad de la EG de diferentes donadores.....	158
1.15	OBJETIVO 3: EFICIENCIA Y ESPECIFICIDAD DE SISTEMAS CRISPR/CAS9 PARA <i>WAS</i> EN CÉLULAS Y HSCs.....	164
1.15.1.1	Eficacia y seguridad de corte en el locus <i>WAS</i>	164
1.15.2	Eficacia y seguridad de RH en el locus <i>WAS</i>	169
	DISCUSIÓN.....	176
	CONCLUSIONES.....	192
	BIBLIOGRAFÍA.....	195

ABREVIATURAS

NOTA

A lo largo del texto se utilizarán abreviaturas de varios términos en inglés. Esto es debido al uso extendido del inglés en publicaciones científicas y al uso de abreviaturas en inglés de manera cotidiana en el lenguaje científico. Para simplificar el texto se mantendrán estas últimas.

A

aa: aminoácidos

ADA-SCID: Síndrome de inmunodeficiencia combinada severa por déficit de adenosina desaminasa

aEJ: NHEJ mediada por microhomologías

B

bGH: hormona de crecimiento bovina

C

c-NHEJ: vía clásica de la NHEJ

CMV: promotor de Citomegalovirus humano

coCas9: secuencia codificante de Cas9 tras la optimización del uso de codones

D

DSB: Roturas de doble cadena en el DNA. Del inglés, double-strand breaks

E

EF1 α : Promotor del factor de elongación 1 α humano

EG: Edición genómica

F

FBS: Suero fetal bovino, del inglés Fetal bovine serum

FISH: hibridación *in situ* fluorescente. Del inglés, Fluorescence *in situ* hybridization

G

GC/ml: Genomas virales por ml. Del inglés, genome copies per ml

H

HDAC: Desacetilación de histonas. Del inglés, histone deacetylation

HDACi: Inhibidor de histona desacetilasas. Del inglés, histone deacetylase inhibitor

hESCs: Células madre embrionarias humanas, del inglés human Embryonic Stem Cells

hESCs: Células madre embrionarias humanas. Del inglés, human Embryonic Stem Cells

HSCs: Células madre hematopoyéticas, del inglés Human hematopoietic stem cells

HOM: células de mucosa oral humana. Del inglés, human oral mucosa

HSF: fibroblastos de piel humana. Del inglés, human skin fibroblasts

I

IDLV: Vector lentiviral deficiente para integración, del inglés

iPSCs: células troncales pluripotentes inducidas. Del inglés, induced pluripotent stem cells

IVIG: Terapia con inmunoglobulinas intravenosas. Del inglés, intravenous immunoglobulin

L

LMC: Leucemia mieloide crónica

LTR: Repeticiones terminales largas. Del inglés, long terminal repeats

LV: Vector lentiviral

M

MFI: Intensidad media de fluorescencia, del inglés mean of the fluorescence intensity

MGNs: Meganucleasas

MLV: Virus de la leucemia murina. Del inglés, murine leukemia viruses.

MPSII: Mucopolisacaridosis tipo II

MSCs: Células mesenquimales estromales, del inglés Mesenchymal stem cells

N

N-WASP: Proteína WASP neuronal

NCs: Células neuronales diferenciadas. Del inglés, neuronal cells

NE: Nucleasas específicas

NHEJ: Unión de extremos no homólogos, del inglés "Nonhomologous end-joining"

NPCs: Células progenitoras neuronales humanas, del inglés Neural Progenitor Cells

P

P/S: Penicilina / Estreptomina

pA: señal de poliadenilación

PAM: Motivo adyacente al protoespaciador

PNAs: ácidos nucleicos peptídicos. Del inglés, peptide nucleic acids

R

RH: Recombinación homóloga

RVD: Diresiduo variable de repetición, de proteínas TALEs. Del inglés, repeat variable diresidue

RVs: Vectores retrovirales

Abreviaturas

SCID-X1: Síndrome de inmunodeficiencia combinada severa ligada al X

S

SCID: Síndrome de inmunodeficiencia combinada severa. Del inglés, Severe combined immunodeficiency

SE: Cassette SFFV-eGFP

T

TFOs: Oligonucleótidos formadores de triplex. Del inglés, triplex forming oligonucleotides

U

UT/ml: Unidades de transducción por ml

W

WAS: Síndrome de Wiskott-Aldrich

WP: Regulador postranscripcional WPRE (del inglés, Woodchuck hepatitis virus posttranscriptional regulatory element)

X

XP: Xeroderma Pigmentoso

Z

ZFNs: Nucleasas con dedos de zinc. Del inglés, zinc-finger nucleases

ZFs: Dedos de zinc, del inglés "Zinc fingers"

γ -RVs: Vectores gammaretrovirales

RESUMEN

Los grandes avances que están teniendo lugar en el campo de la Edición Genómica (EG), particularmente desde la irrupción del sistema CRISPR, están cambiando la forma de abordar la investigación en biología molecular y biomedicina, y plantean un escenario donde estrategias basadas en EG se presenten como alternativa terapéutica para el tratamiento de determinadas enfermedades. Principalmente, podrían verse beneficiadas por este tipo de estrategias ciertas enfermedades monogénicas que hasta la fecha solo cuentan con tratamientos dirigidos a paliar su sintomatología, o para las que acceder a un tratamiento curativo no siempre es posible. En este tipo de afecciones, la reparación del gen alterado mediante EG representa una alternativa prometedora para revertir los síntomas de la enfermedad.

Una de las enfermedades que en un futuro podría ser tratada mediante EG es el Síndrome de Wiskott-Aldrich (WAS). WAS es una inmunodeficiencia primaria, de herencia recesiva ligada al cromosoma X, causada por mutaciones en el gen *WAS*. Este gen codifica para la proteína WASP, específica del linaje hematopoyético, cuyos niveles aparecen ausentes o reducidos en pacientes con WAS. Los afectados muestran inmunodeficiencia celular y humoral, infecciones recurrentes, eczema, microtrombocitopenia con tendencia al sangrado, y propensión al desarrollo de enfermedades autoinmunes y leucemias. La terapia actual para el síndrome de Wiskott-Aldrich incluye el tratamiento de las infecciones, la autoinmunidad, y la trombocitopenia, pero el trasplante alogénico de células madre hematopoyéticas (HSCs) obtenidas de médula ósea es el único tratamiento curativo, hasta la fecha. No obstante, la falta de donantes para algunos pacientes ha contribuido a la búsqueda de terapias alternativas. La modificación de HSCs mediante terapia génica (TG) *ex vivo* es una alternativa para pacientes que no cuentan con un donante compatible. La corrección de la mutación que causa la enfermedad mediante EG sería, en este sentido, la estrategia ideal. Esto se podría conseguir suministrando a las células diana

un DNA donador que albergue la secuencia correcta, y aprovechando la vía de reparación del DNA por recombinación homóloga (RH).

La RH ocurre con baja frecuencia en células somáticas. Sin embargo, el desarrollo en los últimos años de varias tecnologías que incrementan la eficiencia de RH hasta 10.000 veces ha hecho posible su utilización en estrategias de TG. Estas estrategias se basan en el uso de endonucleasas específicas de secuencia que inducen roturas en el DNA, cerca del sitio que se quiere modificar. Estas roturas activan los mecanismos con los que cuenta la célula para reparar el DNA, permitiendo la utilización de un DNA donador como molde para la reparación del daño ocasionado por la endonucleasa.

Dos aspectos fundamentales que deben considerarse para que una estrategia de EG sea trasladable a clínica son eficiencia y seguridad. Una estrategia de EG exitosa debe, por un lado, proporcionar la entrada de las herramientas (nucleasas y DNA donadores) de forma eficiente, en el mayor número posible de células diana, para alcanzar un beneficio terapéutico. Por otro lado, la modificación de las células diana debe llevarse a cabo de una forma segura, sin alterar la biología de las mismas. Conseguir eficiencias de EG altas es relativamente sencillo en líneas celulares inmortalizadas, no así en células primarias, cuya modificación resulta más costosa. En células primarias es, por tanto, necesaria una mejora de los sistemas de entrega de las nucleasas y donadores.

Nuestra hipótesis es que la mejora de la eficiencia y la seguridad de las herramientas de EG permitirán en un futuro cercano la corrección génica en trastornos con base genética, como WAS. Por ello, el objetivo principal de esta tesis consistía en mejorar eficacia y seguridad de las estrategias de EG para edición del locus *WAS*.

En nuestro laboratorio contamos con sistemas CRISPR/Cas9 específicos para edición del locus *WAS*, lo que nos ha permitido comparar, en términos de eficiencia y seguridad, diferentes métodos de entrega para vehicular este sistema a las células

diana: entrega mediante nucleofección del sistema CRISPR/Cas9 en forma de plásmido, o de ribonucleoproteína preformada (RNP), o entrega mediante vectores lentivirales deficientes para integración (IDLVs), con la finalidad de utilizar las propiedades de estos vectores para conseguir niveles óptimos y transitorios del sistema CRISPR/Cas9 en las células diana. Estos análisis se llevaron a cabo en la línea celular K562, de linaje hematopoyético, idónea para modelar WAS, y en nuestras células diana, HSCs y células T. Estas últimas son también diana terapéutica para el tratamiento de WAS, debido a que su modificación resulta más sencilla en comparación con las HSCs, y presentan un menor riesgo de transformación. Además, la expresión de WASP en las células T confiere ventaja selectiva, haciéndolas aún más atractivas.

En K562 observamos que los tres métodos de entrega proporcionaban altos niveles de edición del locus *WAS*, sin que el método de entrega tuviese grandes efectos sobre la especificidad. En HSCs y células T, sin embargo, el único método de entrega que proporcionó niveles de edición de *WAS* robustos fue la nucleofección con RNP. A pesar de los desalentadores datos de la eficacia de los IDLV en células T y HSCs obtenidos inicialmente, seguíamos teniendo la hipótesis de que optimizando el diseño de los IDLVs podríamos lograr los niveles necesarios del sistema CRISPR/Cas9 para editar de forma eficaz y segura estas células. Previamente, en nuestro laboratorio se había diseñado un elemento aislador (IS2), que en vectores lentivirales integrativos (LVs) había mostrado la capacidad de reducir la variabilidad de la expresión y de mejorar los niveles de expresión en células de gran interés. Nos planteamos, sobre la base de estos antecedentes, la inclusión del elemento IS2 en los IDLVs para mejorar su comportamiento. Nuestros datos mostraron que la inclusión del elemento IS2 en el 3'LTR del esqueleto lentiviral mejora los niveles de expresión de los IDLVs a través de un aumento de 6 a 7 veces en la actividad transcripcional de los episomas de IDLVs. Pudimos observar, también, que el comportamiento mejorado de los episomas IDLV-

IS2 probablemente se deba en parte a un reposicionamiento nuclear preferencial en regiones transcripcionalmente activas. Sin embargo, la inclusión de IS2 en el 3'LTR también tenía un efecto negativo en la generación de episomas en las células diana. Por tanto, el efecto final del elemento IS2 en los IDLV depende, en gran medida, de la célula diana y del equilibrio entre los efectos negativos y los efectos positivos en cada tipo de célula. Lamentablemente, en nuestras células diana, células T y HSCs, el efecto neto era nulo o negativo, por lo que estos IDLV-IS2 no se pudieron aplicar a nuestro sistema. A pesar de ello, el efecto neto positivo encontrado en otros tipos celulares de relevancia, como iPSCs, NPCs y neuronas diferenciadas, así como en diferentes líneas inmortalizadas, garantiza la futura aplicabilidad de los IDLV-IS2 en diversas aplicaciones.

Por otro lado, durante el curso de esta tesis se han generado dos modelos celulares reporteros que permiten analizar eficiencia y seguridad de reparación mediante RH de sistemas CRISPR/Cas9 específicos para *WAS*. Estos modelos nos han permitido concluir que las integraciones del DNA donador en sitios indeseados dependen del diseño del DNA donador y del número de dianas en el genoma de la célula. Por otro lado, nos han permitido observar que los DNA donadores que incluyen la diana de corte mejoran la eficacia de RH, sin incrementar la inespecificidad.

Con todos estos antecedentes, en la parte final de la tesis se procedió a abordar la EG del locus *WAS* en células T y HSCs, utilizando los sistemas que mejores resultados nos habían dado en nuestros modelos. Estos datos mostraron que la nucleofección de RNP es el método de elección para la EG de células T y HSCs, y que combinándolo con donadores como productos de PCR, permite editar el locus *WAS* en HSCs y células T. Sin embargo, la eficiencia es todavía mejorable y pudimos detectar varios cortes en sitios fuera de diana, tanto en las células T como en las HSCs. Por lo tanto, aunque la

EG es un abordaje interesante y posible para llevarlo a clínica para el tratamiento de WAS, requiere de nuevas optimizaciones para incrementar la eficacia y la seguridad.

INTRODUCCIÓN

1.1 TERAPIA GÉNICA

La terapia génica (TG) persigue, mediante la introducción de material genético en determinadas células o tejidos, proporcionar un beneficio terapéutico persistente para determinadas enfermedades que actualmente solo cuentan con tratamientos dirigidos a paliar su sintomatología. Resulta particularmente interesante para el tratamiento de enfermedades monogénicas recesivas, donde la inserción del gen terapéutico en el genoma de la célula diana, o la corrección mediante edición genómica (EG) de la mutación causante de la enfermedad, favorecería la restauración en pacientes del fenotipo sano, y su perdurabilidad.

1.2 EDICIÓN GENÓMICA

La EG es una disciplina relativamente novedosa que permite una manipulación precisa de las secuencias del DNA. En sentido amplio, esta modificación dirigida puede producirse a nivel de secuencia, promoviendo la generación de indels (pequeñas inserciones o deleciones), la inserción de una molécula de DNA exógeno en la secuencia diana, o la conversión programada de un par de bases por otro^{1,2}, pero también a nivel epigenético, reprogramando patrones de expresión génica mediante modificación de las marcas de histonas³⁻⁵, así como a nivel postranscripcional, alterando la expresión génica a través de interacciones antisentido con el mRNA^{6,7}. Las técnicas empleadas para la modificación de la secuencia del DNA son las más utilizadas actualmente, y abarcan diferentes estrategias: 1- Aquellas que se basan en la capacidad de los ácidos nucleicos de formar triple hélice, y emplean los oligonucleótidos formadores de triplex (TFOs)^{8,9} y sus derivados, los ácidos nucleicos peptídicos (PNAs)¹⁰⁻¹⁹, 2- Aquellas basadas en el uso de vectores virales adeno-asociados recombinantes (rAAV, del inglés recombinant Adeno-Associated Virus) que favorecen la vía de Recombinación homóloga (RH) dependiente de BRCA-2²⁰, y

3- Aquellas basadas en nucleasas específicas (NE), capaces de generar roturas de forma selectiva en el DNA, promoviendo la modificación de su secuencia principalmente a través de dos mecanismos de reparación del DNA: por unión de extremos no homólogos (NHEJ, del inglés Non-homologous end joining) y mediante RH (Figura 1).

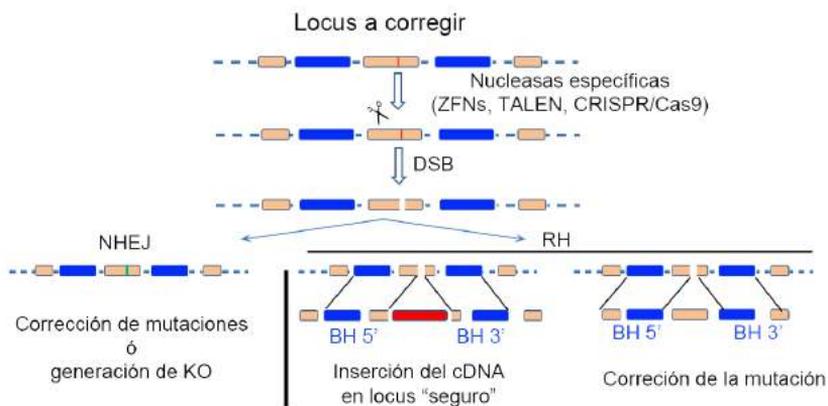


Figura 1. Esquema de los principales mecanismos de edición genómica mediante nucleasas específicas. Se muestran los exones en naranja, y los intrones en azul. La posible mutación a corregir se muestra como una línea roja en mitad del exón. La nucleasa específica cortará cerca del sitio de la mutación causante de la patología, promoviendo que la rotura sea reparada por dos mecanismos principales, según si se suministra o no un DNA donador. A la izquierda se muestra cómo el proceso de NHEJ, mediante la generación de pequeñas inserciones o deleciones, provoca la creación de una nueva mutación o la eliminación de la existente (proceso poco preciso). A la derecha se muestra el proceso de RH, mediado por la presencia de un DNA donador con brazos de homología respecto al sitio de rotura (BH 5' y BH 3'). Dependiendo de la secuencia del donador, se puede restaurar la secuencia mutada, o introducir una secuencia nueva (en rojo) en un locus concreto.

De lo anterior se desprende que las estrategias de EG dirigidas a la modificación de secuencias del DNA se pueden dividir en dos bloques: por un lado, aquellas estrategias que se basan en la utilización de NE y, por otro, aquellas que no usan NE (Figura 2).

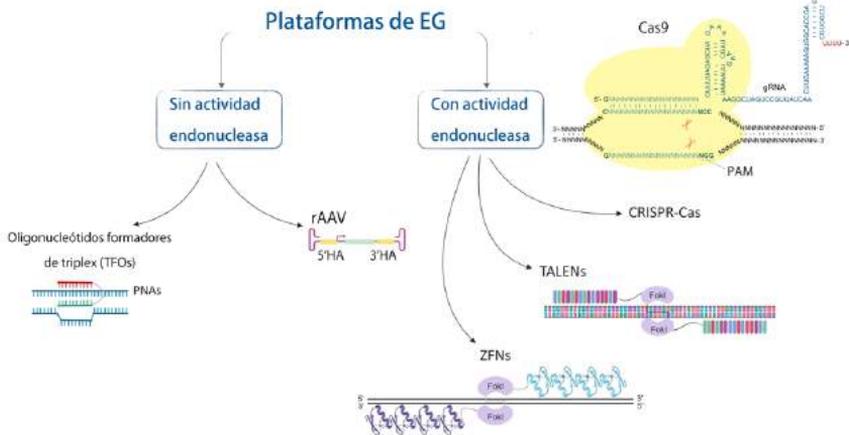


Figura 2. Ilustración mostrando las principales plataformas de EG, las cuales se dividen en dos grupos. En la parte izquierda de la figura se muestra los principales sistemas de EG cuya acción es independiente de actividad endonucleasa. En la parte derecha de la ilustración se muestran los sistemas de edición basados en actividad endonucleasa. rAAV: vector recombinante derivado de virus adeno-asociados. Sistemas CRISPR/Cas: basados en repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (del inglés Clustered Regularly Short Palindromic Repeats). TALENs: nucleasas similares a efectores activadores de la transcripción (del inglés Transcription Activator-Like Effector Nucleases). ZFNs: nucleasas con dedos de zinc (del inglés, Zinc Finger Nucleases).

A continuación se detallan algunos de los sistemas más usados, tanto dentro de las plataformas independientes de endonucleasas, como aquellas que dependen de la actividad endonucleasa.

1.2.1 Plataformas de EG independientes de nucleasa

1.2.1.1 Oligonucleótidos formadores de triplex

Los oligonucleótidos formadores de triplex (TFOs, por sus siglas en inglés) son oligonucleótidos de homopirimidinas, de cadena sencilla, que reconocen en el DNA regiones homopurínicas complementarias, e hibridan con ellas, formando una estructura en triple hélice¹⁴. Los TFOs se acoplan a su diana a través del surco mayor del DNA, y establecen enlaces tipo Hoogsten²¹. La formación de este triplex supone la desestabilización de la hélice de DNA, y la activación de los mecanismos celulares de reparación del DNA, que pueden ser aprovechados para la generación de modificaciones específicas, al suministrar a la célula diana un DNA donador junto al TFO. En esta dirección, existen evidencias de que los TFOs favorecen la RH^{22,23}.

Para mejorar la eficacia de EG de los TFOs, algunos grupos han desarrollado ácidos nucleicos peptídicos (PNAs, del inglés, peptide nucleic acids) que mimetizan las propiedades de los TFOs, pero que presentan mayor estabilidad y afinidad de unión por sus dianas¹⁰⁻¹⁹. Los PNAs, desarrollados por primera vez por Peter E. Nielsen et al. en el año 1993¹⁵, son moléculas con un esqueleto de poliamida, aquirales, y con carga neutra. Son, además, resistentes a nucleasas, y muestran baja citotoxicidad y reducidos efectos *off target*, lo que las convierte en una prometedora herramienta para EG¹⁶⁻¹⁹. Además, sobre los PNAs se pueden introducir diferentes modificaciones para mejorar su eficacia de EG²⁴.

Los PNAs se han empleado exitosamente para EG del locus de la β -globina, asociado con β -talasemia^{24,25}, y del locus CFTR, cuya mutaciones causan fibrosis quística²⁶, así como para regulación antisentido de la expresión de CCR5 en células humanas²⁷, y para su disrupción en HSCs *in vitro*²⁸ e *in vivo*²⁵, en modelo humanizado de ratón.

1.2.1.2 Vectores recombinantes adeno-asociados (rAAV)

Dentro del primer tipo de herramientas de EG, que no requieren la intervención de NE para posibilitar la modificación del DNA, los rAAV han demostrado su capacidad para inducir modificaciones precisas mediante la vía de RH²⁹, tanto *in vitro*¹² como *in vivo*^{20,30}. Esta capacidad de los rAAV de inducir cambios en el DNA de la célula diana preferentemente mediante RH se ha relacionado con la particular estructura de su genoma, consistente en una molécula de DNA de cadena sencilla con un pequeño tramo de doble cadena en la región de las repeticiones terminales invertidas (ITRs, por sus siglas en inglés)²⁹, aunque el mecanismo está aún por dilucidar.

Tanto las plataformas de EG basadas en la formación de triplex, como los rAAV, presentan la ventaja de que no requieren NE exógenas para llevar a cabo la modificación mediante RH del DNA de la célula diana, reduciendo así una posible respuesta inmunitaria³¹ y los efectos genotóxicos, debidos mayoritariamente a las roturas fuera de la diana, y que han sido descritos en algunos casos con el uso de NE.

1.2.2 Plataformas de EG con actividad endonucleasa

Las plataformas de EG basadas en la generación dirigida de roturas en el DNA son, no obstante, las más utilizadas actualmente, tanto en investigación básica, como en aplicaciones clínicas. Entre ellas se encuentran las Meganucleasas (MGNs), las nucleasas con dedos de zinc (ZFNs, del inglés, Zinc Finger Nucleases), las nucleasas de

actividad similar a activadores de la transcripción (TALENs, del inglés Transcription Activator-Like Effector Nucleases) y los sistemas basados en repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (sistemas CRISPR, del inglés Clustered Regularly Short Palindromic Repeats). Estos últimos han supuesto una notable revolución el campo de la EG debido a su simplicidad y versatilidad. Estas herramientas, que comparten la capacidad de generar roturas de doble cadena en el DNA (DSB; del inglés Double Strand Breaks), se denominan de forma global nucleasas específicas de secuencia (NE).

La amplia aceptación de estas plataformas se debe a su capacidad de incrementar la eficiencia de RH más de 1000 veces³² en comparación con sistemas sin actividad endonucleasa. Maria Jasin fue la primera en describir, en 1994, cómo la generación de roturas de doble cadena en el DNA de células de mamífero con la enzima *I-SceI* favorecía la reparación mediante RH, cuando una molécula de DNA con homología respecto al sitio de rotura era suministrada junto a la enzima. Jasin también observó que la reparación de estas roturas podía producirse directamente mediante NHEJ³³. Estos descubrimientos, corroborados por otros grupos poco después³⁴, sentaron las bases sobre las que se fundamentaría el posterior desarrollo de NE para EG.

1.2.2.1 Meganucleasas

Las Meganucleasas (MGNs) son un grupo diverso de endonucleasas que aparecen en la naturaleza, en todos los reinos de la vida, y en múltiples nichos genéticos (incluyendo genomas virales, plásmidos y transposones)^{35,36}. Reconocen secuencias largas de DNA (12-45pb), muy poco frecuentes. En relación con sus motivos estructurales, hasta el momento se han descrito seis familias de MGNs³⁷, siendo la más estudiada y mejor caracterizada la familia de proteínas LAGLIDADG. Las proteínas

LAGLIDADG, consideradas parásitos moleculares y codificadas generalmente en intrones móviles o inteínas, presentan actividad endonucleasa muy específica, que facilita la movilización de su secuencia codificante, y/o actividad madurasa, relacionada con su autosplicing. El motivo LAGLIDADG desempeña un papel tanto a nivel estructural, como catalítico³⁸. El reconocimiento de la secuencia de DNA radica en una serie de interacciones DNA-proteína que se establecen entre la secuencia diana y láminas β del dominio endonucleasa, acopladas a través del surco mayor del DNA, y que comprenden enlaces de hidrógeno e interacciones estéricas que refuerzan la especificidad³⁹.

La aplicabilidad de las MGNs como herramientas de EG se ve reducida por el limitado número de dianas naturales que reconocen. Tratando de ampliar el repertorio de MGNs, algunos grupos han alterado la especificidad de determinadas MGNs mediante evolución dirigida, mutando residuos específicos⁴⁰⁻⁴⁵. También se han desarrollado MGNs quiméricas, intercambiando diferentes dominios⁴⁶⁻⁴⁹.

Las MGNs artificiales han sido usadas para restaurar mutaciones en los genes humanos XPC y RAG1, entre otros, relacionadas respectivamente con el Xeroderma Pigmentoso (XP) y con el Síndrome de inmunodeficiencia combinada severa (SCID)^{50,51}. Sin embargo, la dificultad de reorientar la especificidad de reconocimiento de las MGNs, junto al gran desarrollo que están alcanzando NE alternativas, más fácilmente personalizables, han desplazado a las MGNs como herramientas de EG.

1.2.2.2 ZFNs

La función de reconocimiento del DNA y la función catalítica están estrechamente ligadas en la mayoría de las enzimas de restricción, de forma que la introducción de mutaciones en el dominio de unión al DNA podría alterar la función catalítica. No así en el caso de *FokI*, endonucleasa de restricción del procarionta *Flavobacterium okeanoikoites*⁵², donde aparece una estructura bimodular, con un dominio N-terminal de unión al DNA, y un dominio C-terminal con actividad endonucleasa no específica, separados. Sobre la base de esta estructura, Chandrasegaran y colaboradores exploraron la posibilidad de intercambiar el dominio original de reconocimiento del DNA de *FokI* por otras proteínas de unión a DNA, para redirigir la especificidad de la endonucleasa⁵³. Así, el equipo de Chandrasegaran desarrolló la primera endonucleasa quimérica en 1994, mediante la adhesión del dominio catalítico genérico *FokI* al homeodominio Ubx de *Drosophila*, proteína con motivo hélice-bucle-hélice de unión a secuencias conservadas presentes en varios genes de *Drosophila*⁵⁴. Una vez probada su efectividad, fueron ellos mismos quienes dos años más tarde desarrollaron las primeras ZFNs a partir de la fusión del dominio endonucleasa *FokI* a proteínas con dedos de zinc Cys₂His₂⁵⁵.

Las proteínas con dedos de zinc Cys₂His₂ están compuestas por varios dominios de dedos de zinc (ZF, del inglés Zinc Finger) en tándem. Los dedos de zinc Cys₂His₂ representan uno de los motivos estructurales más frecuentes en eucariotas, presentes en proteínas implicadas en la regulación de la expresión génica, como sucede en aproximadamente el 45% de factores de transcripción humanos. Se caracterizan por ser dominios de unión al DNA compuestos por dos láminas β antiparalelas, y una α -hélice. Cada dominio ZF reconoce 3-4 nucleótidos de forma específica, siguiendo un código de reconocimiento determinado por cuatro amino ácidos, en posiciones específicas del extremo N-terminal de la región α -hélice, encargados de interactuar

con el DNA a través de su surco mayor ⁵⁶⁻⁵⁹. Otras interacciones entre residuos en posiciones no canónicas y el DNA, que incluyen puentes de hidrógeno e interacciones no covalentes, como fuerzas de van der Waals, también contribuyen a esta especificidad ⁶⁰.

La síntesis de dominios artificiales de unión al DNA con nuevas especificidades es posible mediante la combinación de 3-6 dominios ZF. La fusión de estos dominios al dominio catalítico *FokI* conforma un monómero ZFN, con la capacidad de reconocer una secuencia de DNA de 9 a 18pb de longitud⁶¹ y conducir, así, la actividad endonucleasa a la diana deseada. El dominio catalítico de *FokI* debe dimerizar, y requiere dos copias invertidas del sitio de unión al DNA (una en cada una de las hebras), para generar de forma eficiente DSBs⁶²⁻⁶⁴. De esta forma, dos monómeros ZFN, cuyas dianas se encuentran en cadenas opuestas, conforman la arquitectura canónica de las ZFNs (Figura 3).

La especificidad de esta arquitectura se mejoró mediante la modificación del dominio *FokI* para prevenir la homodimerización y supeditar la creación de DSBs a la formación de heterodímeros^{65,66}. Recientemente, Paschon et al. han desarrollado ZFNs capaces de adoptar configuraciones diméricas alternativas a la arquitectura canónica, aumentando así las opciones de diseño y la precisión de las ZFNs⁶⁷.

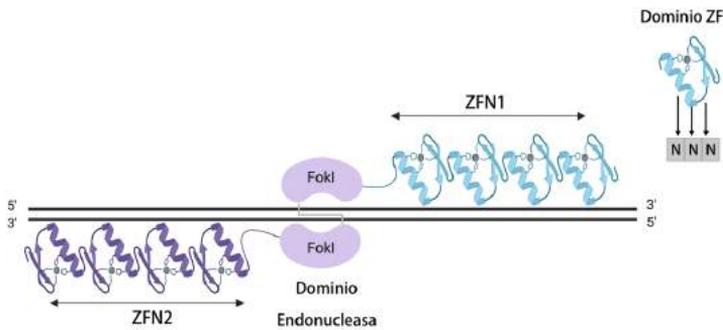


Figura 3. Ilustración de la arquitectura canónica de las ZFNs. Esquema de un heterodímero de ZFNs reconociendo a su diana. Cada ZFN contiene el dominio de corte *FokI* adherido a un array de tres a seis dominios ZF (cuatro en este caso), diseñados para reconocer específicamente la secuencia que flanquea al sitio de corte. Cada dominio ZF reconoce 3-4 nucleótidos de forma específica.

Desde su creación, las ZFNs han sido una de las piezas clave en el avance de la EG. Tras demostrar su capacidad para estimular la RH⁶⁸⁻⁷⁰, fueron las primeras NE empleadas para modificar por RH un locus endógeno en células primarias humanas (en concreto, en células T CD4⁺)⁷¹. Las ZFNs permitieron, también por primera vez, la modificación de organismos completos mediante EG^{72,73}. Si bien, el mayor éxito de las ZFNs es el potencial terapéutico que han demostrado, llegando a ser las herramientas empleadas en el primer ensayo clínico de EG, para la disrupción del correceptor CCR5 en células T autólogas, para conferir resistencia a la infección por VIH⁷⁴. Así mismo, están siendo empleadas en ensayos clínicos más recientes, por ejemplo, para la restauración de la eritropoyesis en β -Talasemia (clinicaltrials.gov; NCT03432364)⁷⁵, para el tratamiento de Hemofilia B, mediante la inserción en hepatocitos de una copia correcta del Factor IX de coagulación en el locus de la albúmina (clinicaltrials.gov; NCT02695160), así como para el primer ensayo clínico de EG *in vivo*, aprobado para el tratamiento de Mucopolisacaridosis tipo II (MPSII) (clinicaltrials.gov; NCT03041324).

1.2.2.3 TALENs

Las nucleasas de tipo efectores activadores de la transcripción (o TALENs, por sus siglas en inglés) son unas endonucleasas de carácter modular, compuestas en su extremo C-terminal por el dominio catalítico *FokI*, asociado a un dominio de unión secuencia-específica a DNA. Este segundo dominio tiene su origen en la familia de proteínas TALEs, secretadas por bacterias fitopatógenas del género *Xanthomonas*, y que actúan como activadores de la transcripción en el núcleo de las células vegetales que infectan estas bacterias⁷⁶. El dominio TALE se compone por repeticiones en tándem de 33-35 aminoácidos (aa) altamente conservados, a excepción de los aa 12 y 13, denominados de forma conjunta diresiduo variable de repetición (RVD, por sus siglas en inglés), cuya función es el reconocimiento de un par de bases concreto⁷⁷⁻⁷⁹.

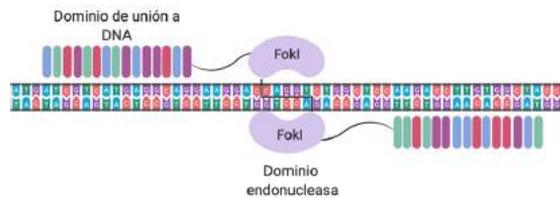


Figura 4. Estructura de las TALENs. El dominio de unión al DNA contiene repeticiones en tándem de 33-34 aa altamente conservados, a excepción de los aminoácidos 12 y 13 hipervariables (RVD), que dictan la especificidad de unión al DNA. En las TALENs, el dominio de unión al DNA se fusiona a un dominio efector, consistente en la endonucleasa *FokI*. Esquema representativo de la unión de las TALENs al DNA diana y de los sitios de corte de las nucleasas *FokI*.

1.2.2.4 CRISPR/Cas

La tecnología CRISPR, basada en una endonucleasa específica cuya actividad es dirigida por una pequeña molécula de RNA, ha sido la última en incorporarse al abanico de herramientas de EG, y ha supuesto una gran revolución en este campo, por su flexibilidad y fácil diseño.

Los sistemas CRISPR/Cas derivan de un sistema inmune adaptativo propio de bacterias y arqueas, cuyos genomas contienen una serie de repeticiones cortas palindrómicas (23-50pb), interespaciadas por secuencias de origen viral o plasmídico (17-84pb), denominadas secuencias espaciadoras, adquiridas durante infecciones pasadas (locus CRISPR, por sus siglas en inglés). Al locus CRISPR precede una secuencia líder rica en A-T, y próximos a él aparecen un conjunto de genes, denominados genes cas, que codifican las proteínas Cas que, junto al locus CRISPR, actúan como maquinaria de defensa frente a bacteriófagos y plásmidos⁸⁰. A partir del locus CRISPR se generan transcritos primarios de cadena larga, que son posteriormente procesados en un conjunto de RNAs cortos (crRNAs), que contienen una repetición conservada y una secuencia espaciadora variable, complementaria al ácido nucleico invasor. La combinación de estos crRNAs con proteínas Cas forma un complejo efector, que previene la proliferación y propagación de los elementos genéticos foráneos, mediante la generación de roturas específicas de secuencia⁸¹.

Esta serie de repeticiones palindrómicas interespaciadas se detectaron, por primera vez en 1987, en *Escherichia coli*⁸². Sin embargo, la naturaleza de estas secuencias no fue descifrada hasta algo más de una década después⁸³⁻⁸⁵. En base a la filogenia de los genes cas, la secuencia y organización de las repeticiones del locus CRISPR, así como en base a su arquitectura, se han definido tres tipos diferentes de sistemas CRISPR/Cas^{86,87}. El primer sistema CRISPR/Cas en ser adaptado para EG en células

eucariotas, y el más ampliamente extendido, es el sistema tipo II CRISPR/Cas9 de *Streptococcus pyogenes*⁸⁸. Este sistema está formado por la proteína Cas9, de 1368 a.ác⁸⁹, y contiene dos dominios nucleasa RuvC cerca del extremo amino terminal y un dominio HNH en el centro de la proteína. A su vez, Cas9 forma un complejo con dos moléculas de RNA, crRNA y tracrRNA, que juntas forman el denominado RNA guía (gRNA, por sus siglas en inglés). El híbrido crRNA-tracrRNA guía a la proteína Cas9 para escindir el DNA diana mediante la generación de DSBs. Próximo al protoespaciador se encuentra el “motivo adyacente al protoespaciador” o motivo PAM, muy conservado, y aguas arriba del sitio de unión del crRNA. Esta combinación de varias moléculas hace del sistema CRISPR/Cas un sistema útil, que puede ser rediseñado de forma sencilla, con la única modificación del gRNA, para así, redirigirlo hacia otra diana. Por ello, actualmente el sistema CRISPR/Cas se está utilizando con éxito como herramienta de edición génica. Posteriormente, se han desarrollado varios sistemas CRISPR/Cas basado en diferentes especies bacterianas.

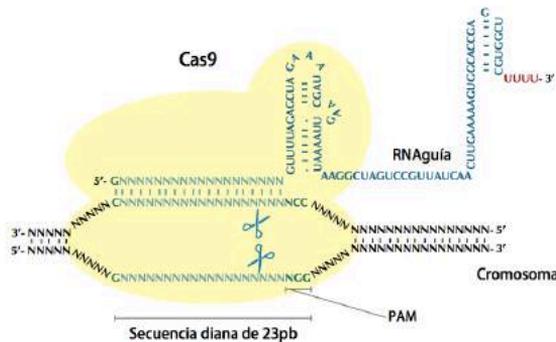


Figura 5. Esquema representativo del Sistema CRISPR/Cas9. Requiere la co-expresión de la proteína Cas9, y uno o varios gRNA. Cas9 desarrollará el dúplex de DNA y se unirá a su secuencia diana, pero solo si esta contiene la secuencia PAM correcta en 3'.

1.3 MECANISMOS DE REPARACIÓN DEL DNA

La preservación de la integridad genómica es crucial para la supervivencia de células eucariotas, por lo que disponen de diferentes mecanismos para reparar las distintas lesiones que se puedan producir en su DNA, incluyendo las DSBs, que constituyen la forma de daño en el DNA más citotóxica⁹⁰. La no reparación de una DSB produce la entrada de la célula en apoptosis o senescencia, mientras que una reparación inapropiada puede conducir a inestabilidad genómica y desarrollo tumoral. Las células eucariotas cuentan con dos mecanismos principales de reparación de DSBs, que difieren entre sí en la fidelidad de la reparación y en la necesidad de un molde para la misma. Estos dos mecanismos son, la unión de extremos no homólogos, o NHEJ, y la RH. Las estrategias de EG basadas en la creación de DSBs mediante NE se sirven de estos mecanismos de reparación para introducir en el genoma de las células diana las modificaciones deseadas, aunque también se pueden utilizar otras vías (Figura 6).

1.3.1 NHEJ

La NHEJ promueve la adición y/o eliminación de nucleótidos en los extremos generados por una DSB, y su religación para completar el proceso de reparación, y es, en todas las células de mamífero, la vía de reparación del DNA predominante^{91,92}. La NHEJ se da en todas las fases del ciclo celular, aunque la determinación de su contribución a la reparación de DSBs durante las distintas fases depende de los modelos sobre los que se analice, y los mecanismos empleados para ello⁹³⁻⁹⁵. Además, puede ocurrir por al menos dos vías distintas: una vía dependiente del heterodímero Ku, de XRCC4 y de ligasa IV, o vía clásica (c-NHEJ), y otra alternativa, mediada por microhomologías (aEJ)^{91,96} (Figura 6).

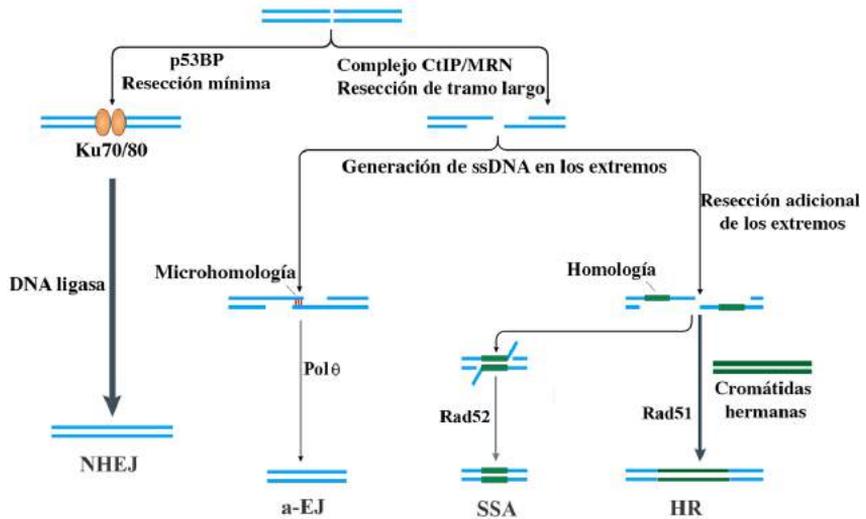


Figura 6. Figura que muestra la elección del mecanismo de reparación de DNA. Las DSBs son corregidas, principalmente, mediante NHEJ o mediante RH, aunque existen mecanismos alternativos como son la unión de extremos alternativa (a-EJ, del inglés alternative end joining) y la unión de cadena simple (SSA, del inglés single strand annealing) (adaptado de Chang et al., 2017⁹⁷)

En células somáticas humanas se da principalmente la c-NHEJ⁹⁸, en la que se pueden distinguir cuatro fases⁹⁹: **1.** Reconocimiento de la rotura, ensamblaje y estabilización del complejo de NHEJ en la DSB. **2.** Establecimiento de un puente entre ambos extremos. **3.** Procesamiento de los extremos. **4.** Ligación de los extremos y disolución del complejo de NHEJ. Así, de forma resumida, la c-NHEJ se inicia por la unión del heterodímero Ku (compuesto por las subunidades Ku70 y Ku80, de 70 y 80 kDa respectivamente) a ambos extremos de la DSB, unión que se produce de forma inmediata tras la generación de la DSB. Ku parece funcionar atrayendo hacia los extremos de la rotura a la DNA-PKcs, (subunidad catalítica de la proteína quinasa dependiente de DNA) para la formación de un complejo DNA-Ku-DNA-PKcs, que actúa

como puente entre ambos extremos de la rotura¹⁰⁰. Ku y DNA-PKcs actúan reclutando y estabilizando proteínas de procesamiento encargadas de convertir extremos no compatibles en extremos religables (como la proteína Artemis, con actividad endo y exonucleasa, encargada del procesamiento de extremos 5' y 3' protuberantes¹⁰¹, o polimerasas), así como las proteínas XRCC4, XLF y DNA ligasa IV, que formarán el complejo XRCC4/XLF/Ligasa IV, encargado de ligar ambos extremos^{102,103}.

Dado que este sistema de corrección no utiliza la secuencia homóloga para restaurar las DSBs, se producen con frecuencia errores, consistentes en deleciones o inserciones de nucleótidos (indels)¹⁰⁴. El mecanismo de c-NHEJ puede ser aprovechado, por ejemplo, para restaurar el marco de lectura de una proteína truncada^{105,106}, o para la disrupción de un locus mediante EG^{74,107,108}.

En lo que se refiere a la resolución mediante NHEJ de DSBs generadas por el sistema CRISPR/Cas9, se ha descrito un patrón no aleatorio del tipo y tamaño de indels generados. Recientemente, Chakrabarti et al., sobre una batería de 1491 dianas y sus correspondientes sgRNAs, observaron que este patrón depende considerablemente del sitio diana, y que el tipo de indel más frecuente lo constituyen inserciones o deleciones de un solo nucleótido. También observaron una correlación entre la eficiencia de EG y la precisión, de forma que en aquellas dianas con un patrón definido de indels era en las que se obtenía una mayor eficiencia, así como una relación entre la secuencia diana y la precisión de EG, y entre el estado de la cromatina y la formación de indels. Observaron una influencia de las posiciones 5, 4, 3, 2 aguas arriba de la secuencia PAM (núcleo de precisión de la secuencia diana) sobre la precisión de EG, siendo la posición 4, que además coincide con la posición en 5' del sitio de corte, la más determinante¹⁰⁹.

1.3.2 Recombinación Homóloga

La RH permite el intercambio de DNA entre secuencias idénticas o similares, y actúa favoreciendo la variabilidad genética durante la meiosis, así como durante la generación del repertorio inmunológico, en la recombinación V(D)J de linfocitos T y B, pero también tiene un papel fundamental en la reparación de DSBs y otras lesiones del DNA. A diferencia de la NHEJ, la RH se restringe a las fases S y G2 del ciclo celular^{93,94,110}, ya que requiere para la reparación la actuación como molde de la cromátida hermana de aquella en la que se ha producido el daño¹¹¹, lo que hace de la RH una reparación conservativa^{112,113}.

De forma muy resumida, la RH comienza con el reconocimiento del daño y posterior reclutamiento de proteínas de señalización de daño y de reparación del DNA. El complejo MRN, compuesto por MRE11/NBS1/RAD50, actúa como sensor de daño, y es uno de los primeros en llegar a la rotura¹¹⁴. Este complejo activa, a continuación, a las quinasas ATM y ATR, encargadas de señalar la detención del ciclo celular para permitir la reparación del DNA¹¹⁴. Los extremos se procesan y se eliminan nucleótidos por actividad exonucleasa en dirección $5' \rightarrow 3'$, lo que genera un fragmento de cadena sencilla que contiene el extremo 3'OH libre^{115,116}. El extremo generado es recubierto por la proteína RPA, que protege el DNA contra la degradación y la formación de estructuras secundarias¹¹⁷, y favorece el reclutamiento de diversas proteínas del complejo RAD51. Este complejo inicia la búsqueda de homología entre la región dañada y el DNA molde, formando un filamento nucleoprotéico capaz de invadir y desplazar una de las hebras de la molécula no dañada, y de formar un heterodúplex de DNA entre la región complementaria de ambas. A partir del extremo 3'OH del sitio original de rotura tiene lugar la síntesis de DNA en el heterodúplex, donde la hebra complementaria proporciona un molde ideal para que la reparación tenga lugar libre de errores. Finalmente, el complejo Resolvasa, dependiente de RAD51C-XRCC3¹¹⁸, es

el encargado de resolver la RH. La hebra de nueva síntesis es liberada, y aparea ahora con su cadena complementaria original.

1.4 LIMITACIONES DE LA EG: EFECTOS *OFF TARGET*

A medida las técnicas de EG han ido avanzando hacia un uso clínico de las células editadas genéticamente, crece la preocupación sobre la especificidad de estas herramientas. En otras palabras, la aplicación clínica de las células editadas genéticamente debe tener unos estándares de seguridad elevados, lo que depende claramente de una determinación fiable de la especificidad de estas herramientas. Detectar, y en su caso reducir o eliminar estos cortes fuera de diana, es imprescindible.

En las siguientes líneas se van a discutir los diferentes métodos que se han ido desarrollando para identificar y caracterizar los sitios fuera de diana, enfocándonos en el sistema CRISPR/Cas9, debido al gran interés que está despertando este sistema CRISPR de Clase II y sus variantes, y dado que en el desarrollo de esta memoria nos hemos centrado fundamentalmente en él. No obstante, muchas de las metodologías empleadas para evaluar la actividad en sitios *off target* son aplicables a otras nucleasas (por ejemplo, ZFNs, TALENs, etc).

Las NE pueden generar DSBs en regiones del DNA fuera del sitio diana (sitios *off target*, por su nomenclatura en inglés), con un cierto grado de homología respecto a este. La generación de DSBs en sitios *off target* puede producir inserciones o deleciones (indels), así como translocaciones (si ocurre más de un corte simultáneamente), que podrían ser perjudiciales para la célula. Por otro lado, cuando una estrategia de EG está diseñada para modificar un locus concreto mediante RH, también se puede producir la inserción del donador en sitios *off target* en los que la endonucleasa esté

produciendo cortes de forma inespecífica, así como en sitios en los que se produzcan roturas espontáneas del DNA.

1.4.1 Métodos sesgados de detección de sitios *off target*

Las primeras estrategias empleadas para identificar actividad de los sistemas CRISPR/Cas9 en sitios *off target* se basaron en la predicción *in silico* de posibles sitios *off target*, y posterior análisis de la formación de heterodúplex en estos sitios, o posterior secuenciación de los mismos. Aunque la secuenciación masiva es el método de referencia para medir las frecuencias de mutagénesis fuera de sitio¹¹⁹⁻¹²¹, una serie de métodos alternativos como los algoritmos TIDE y ICE^{122,123}, que identifican las principales mutaciones en un sitio concreto mediante técnicas de descomposición de secuencias Sanger en muestras procedentes de una población celular, o las técnicas empleadas con menos frecuencia, basadas en las enzimas T7 endonucleasa (T7E1) y CEL I, nucleasas muy sensibles al desajuste que hay entre hebras de DNA. Los algoritmos TIDE y ICE estiman la frecuencia indels dentro de una población de alelos editados mediante la descomposición de los cromatogramas de Sanger. En comparación con las técnicas basadas en el uso de T7 y CEL I, estos algoritmos proporcionan una estimación más precisa de la fracción del genoma editado dentro de una población mixta, así como su identidad. Lógicamente, esta metodología basada en la descomposición de secuencias está limitada por la calidad de los cromatogramas¹²⁴.

Los métodos de detección mencionados previamente se basan generalmente en algoritmos de predicción que pronostican una serie de potenciales sitios *off target*, que presentan cierta similitud con la secuencia diana, considerando la secuencia del gRNA, así como de la secuencia (tanto canónica, como no canónica) de la región PAM (Tabla 1). Algunos algoritmos simplemente califican los sitios *off target* mediante

análisis de similitud de secuencia. Sin embargo, otros utilizan un cálculo más complejo, que tiene en cuenta tanto la secuencia, como otros parámetros de calidad. En este sentido, recientemente Doench et al. han desarrollado una puntuación que tiene en cuenta, no solo la diferencia en la secuencia, sino también la posición de esos desapareamientos en la secuencia del guía¹²⁵.

Los sistemas basados en predicciones *in silico*, aunque son métodos potentes y útiles, no predicen todos los potenciales sitios *off target*. Recientemente, Tsai et al, así como Cameron et al., demostraron que los sistemas de predicción *in silico*, como E-Crisp, CCTop y Cas-OFFinder (Tabla1), son capaces de detectar solamente parte de los *off target* detectados experimentalmente. Estos trabajos pusieron énfasis sobre la necesidad de desarrollo de sistemas de predicción mucho más certeros.

Tabla 1. Sistemas de predicción *in silico* de sitios *off target*

Nombre de la plataforma	Sitio Web
E-CRISP	http://www.e-crisp.org/E-CRISP/
CCTop	https://crispr.cos.uni-heidelberg.de/
Cas-OFFinder	http://www.rgenome.net/cas-offinder/
CRISPRseek	https://bioconductor.org/packages/release/bioc/html/CRISPRseek.html
COSMID	https://crispr.bme.gatech.edu/
CROP-IT	http://www.adlilab.org/CROP-IT/homepage.html
CasFinder	http://arep.med.harvard.edu/CasFinder/

Fu et al.¹²⁶ publicaron el primer estudio para señalar que el sistema CRISPR / Cas9 puede generar altos niveles de corte en sitios *off target* bajo ciertas condiciones. Desde entonces, numerosos grupos de investigación han producido abundantes datos en sitios *off target* utilizando diferentes estrategias¹²⁷.

No obstante, tal y como se ha comentado, es muy complicado diseñar un algoritmo que, para la predicción de sitios *off target*, tenga en consideración ciertos factores, como el contexto genómico. Esto supone que los algoritmos empleados para la predicción *in silico* no son capaces de predecir todos los posibles sitios *off target* para un gRNA determinado, por lo que es necesario recurrir a métodos no sesgados para determinar la bioseguridad de una estrategia de EG.

1.4.2 Métodos no sesgados de detección de *off target*

Para sortear las limitaciones asociadas a los métodos predictivos *in silico*, se han desarrollado métodos no sesgados que permiten una visión global de todos los sitios *off target*. Los métodos de detección no sesgados permiten la identificación de sitios en los que se ha producido cortes en células diana sin tener en cuenta ningún tipo de predicción. Estos métodos no sesgados se dividen en dos grupos fundamentales: métodos *in cellula* o métodos *in vitro*. El primer grupo, utiliza capturas directas de células que han sido editadas mediante un sistema de edición concreto. Se basan, fundamentalmente, en la captura de una molécula de DNA proporcionada exógenamente, y en la lectura de translocaciones cromosómicas. Más en desuso, existen técnicas basadas en sistemas libres de células, utilizando DNA genómico purificado.

En la

Tabla 2 se citan los métodos no sesgados más utilizados hasta la fecha.

Tabla 2. Métodos no sesgados para la identificación de sitios *off target*

Nomenclatura	Referencia
Captura de IDLVs	126
GUIDE-seq	127
High-throughput, genome-wide translocation sequencing (HTGTS)	128
BLESS	129
BLISS	130
ChIP-dCas9	131

A continuación, se resume la base de tres de estos métodos:

GUIDE-seq: este método, desarrollado en el laboratorio de Young¹³², se basa en la incorporación de un un fragmento de DNA marcado, un oligodesoxinucleótido de doble cadena (dsODN, por sus siglas en inglés), al sitio de corte. Una vez insertado el dsODN, la secuencia adyacente al sitio de corte es amplificada selectivamente, y sometida a secuenciación masiva. Quizás la mayor limitación de la técnica basada en GUIDE-Seq es la necesidad de una alta eficiencia de transfección de los dsODN que sirven de marcaje. A pesar de esta limitación, este método ha sido utilizado para evaluar la especificidad de varios sistemas de edición^{127,133,134}.

Inmunoprecipitación de cromatina ChIP-dCas9: esta estrategia persigue detectar todos los sitios de unión de un determinado complejo gRNA-Cas9. Es un método indirecto que parte de la premisa de que la presencia de un complejo gRNA-Cas9 en un locus siempre generará un DSB, y para garantizar la unión estable del complejo gRNA-Cas9 al locus objetivo, emplea una versión catalíticamente inactiva de Cas9 (dCas9). Mediante ChIP-seq, que combina la inmunoprecipitación de cromatina con secuenciación de segunda generación, se produce una lectura clara de los sitios de unión del complejo dCas9-gRNA, tanto en la diana (*in target*), como en localizaciones alternativas (sitios *off target*)¹³⁵⁻¹³⁸.

Captura de Vectores lentivirales no integrativos (IDLVs): Este fue el primer método no sesgado empleado para determinar efectos *off target* de ZFNs, y ha sido usado también con otras nucleasas específicas, incluyendo el sistema CRISPR/Cas9. Los IDLVs permanecen como episomas en el núcleo de las células en que entran, pero pueden integrarse en sitios en los que se producen DSBs. La principal ventaja de esta estrategia es la alta eficiencia con que IDLVs entran en el núcleo de importantes células diana, incluyendo células primarias difíciles de transfectar.

1.5 EDICIÓN GENÓMICA EN TERAPIA GÉNICA

Durante la última década, la TG clásica, basada en la adición de genes ha logrado éxitos sin precedentes, logrando la aprobación de 7 medicamentos de TG para enfermedades hereditarias (Glybera, Luxturna, Strimvelis) y cáncer (Imlygic, Gendicine, Kymria) y varios estudios clínicos en etapas avanzada de desarrollo que prometen nuevos medicamentos próximamente. Sin embargo, a pesar de estos éxitos, todavía hay varios aspectos de estas tecnologías que deben mejorarse para aplicaciones amplias en la clínica. Uno de estos aspectos se refiere a la forma en que las células diana se modifican genéticamente, que involucran mantenimiento episomal o integración semi-aleatoria de los transgenes terapéuticos. Estas formas clásicas pueden causar pérdida de beneficios terapéuticos con el tiempo y / o posibles efectos secundarios debido a integraciones de transgenes en genes clave.

La EG ha abierto un amplio panel de nuevas posibilidades, no solo para mejorar las estrategias de adición genética, sino también para el desarrollo de otras nuevas que son potencialmente más eficientes y más seguras. Las ventajas más importantes son la posibilidad de reparar la mutación causante de la enfermedad y elegir la ubicación genómica desde donde expresar el transgén terapéutico (puerto seguro). Sin embargo, hay varias limitaciones para la aplicación clínica de la EG, tales como las incertidumbres con respecto a la actividad *off-target* de las NE, la alta proporción de reparación mediante NHEJ en comparación con la RH, que conduce a mutaciones no deseadas, así como a las dificultades para entregar las herramientas de EG en algunas células diana importantes como neuronas, células musculares y HSCs.

A pesar de las limitaciones mencionadas anteriormente, ya hay 23 ensayos clínicos de TG en curso basados en GE que utilizan ZFNs (14), CRISPR / Cas9 (16) y TALENs (3) para el tratamiento de enfermedades infecciosas (VIH-1, VPH), cáncer, trastornos

sanguíneos y trastornos metabólicos (Clinicaltrials.gov, mayo de 2019). Estos éxitos muestran el potencial de la EG como estrategia para TG y han venido de la mano de las mejoras en los vectores de transferencia génica que ahora han alcanzado eficiencias suficientes en varias células humanas primarias.

1.5.1 Vectores para terapia génica

Una de las claves que determinará el éxito de cualquier estrategia de TG, incluidas las de EG, consiste en la elección del sistema de entrega que hará llegar al núcleo de las células diana las distintas herramientas requeridas (nucleasas y donadores, en el caso de estrategias de EG). El método de entrega escogido debe superar las barreras fisiológicas, protegiendo la integridad del material genético y garantizando la adecuada internalización y posterior liberación de este material dentro de la célula diana, enfrentando una mayor complejidad en estrategias *in vivo*. Por ello, al escoger el método de entrega es muy importante considerar factores como el tipo celular, el microambiente celular, y las potenciales respuestas inmunitarias.

En base a su naturaleza, los vectores de transferencia se pueden clasificar en vectores no virales, y vectores virales.

La capacidad inherente de los virus de transportar material genético al interior de las células que infectan los convierte en vehículos de transferencia génica ideales. Aprovechar esta característica requiere, no obstante, que se introduzcan modificaciones en el genoma viral de forma que sean deficientes para replicación en la célula diana.

1.5.1.1 Vectores Adenovirales (AdV)

Los AdV son virus sin envuelta, con cápside protéica icosaédrica, cuyo genoma es una molécula de DNA bicatenario lineal de longitud en torno a 26 – 40 Kb¹³⁹, flanqueado a ambos lados por repeticiones terminales invertidas (ITRs, del inglés Inverted Terminal Repeats). Son capaces de infectar tanto células en división como células quiescentes, permaneciendo su genoma de forma episomal en la célula infectada. Hasta la fecha se han descrito más de 50 serotipos de AdV humanos, así como numerosos serotipos de AdV de otras especies.

La gran capacidad de empaquetamiento de los AdV, así como su alta eficiencia a la hora de transducir diferentes tipos celulares, ha propiciado que se destinen numerosos esfuerzos a la optimización de AdV como vectores de transferencia génica. Sin embargo, el uso de AdV como vectores se ve comprometido por su alta inmunogenicidad¹⁴⁰. Los vectores adenovirales (rAdV) de primera generación desencadenaban respuestas inmunitarias muy fuertes, debido a que la mayoría de personas han sido infectadas con AdV y posee anticuerpos neutralizantes, al menos frente a los serotipos más comunes de AdV. La última versión de rAdV presenta varias ventajas con respecto a los rAdV de primera generación; solo conservan las secuencias víricas correspondientes a los ITRs y la señal de encapsidación, y tienen gran capacidad de empaquetamiento (hasta 35Kb), aunque las proteínas de la cápside aún podrían desencadenar una respuesta inmunitaria.

A pesar de esta limitación, los rAdV han sido ampliamente usados como vectores oncolíticos, así como para TG, y para el desarrollo de vacunas expresando antígenos foráneos, llegando a numerosos ensayos clínicos¹⁴¹. En EG, Perez et al. emplearon los rAdV como sistema de entrega en el primer estudio que puso de manifiesto que la disrupción de CCR5 en células T humanas mediante ZFNs suprimía la replicación del

HIV-1¹⁰⁷, y que posteriormente daría lugar al primer ensayo clínico de EG en fase I, dirigido por Pablo Tebas y Carl June⁷⁴. En el caso del Sistema CRISPR, los rAdV han sido empleados por varios grupos para la entrega de Cas9, tanto para el establecimiento de modelos de enfermedad, como para la investigación de tratamientos. Por ejemplo, también se han usado para generar linfocitos T CD4⁺ resistentes a la infección por VIH mediante la disrupción de CCR5¹⁴², o para restaurar la expresión de la distrofina en mioblastos derivados de paciente con distrofia muscular de Duchenne (DMD)¹⁴³. También han sido usados para la entrega de Cas9 *in vivo*, por ejemplo, para desarrollar modelos de cáncer de pulmón en ratón mediante la generación de reordenamientos cromosómicos¹⁴⁴, para editar distintos genes de forma específica en hígado de ratón^{145,146}, para restaurar la expresión de la distrofina en ratones mdx¹⁴⁷, o para revertir la hemofilia B en un modelo de ratón para hemofilia, mediante la inserción dirigida del cDNA del factor IX de coagulación en el sitio seguro ROSA26¹⁴⁸. En este último caso, tanto la entrega del sistema CRISPR/Cas9, como la del donador codificando el factor IX de coagulación, se realizaron mediante rAdV. No obstante, en algunos de estos estudios se han descrito respuestas inmunitarias frente al vector adenoviral, así como frente a Cas9^{145,148}.

1.5.1.2 Vectores Adeno-Asociados (rAAV)

Entre los vectores de origen viral destacan los vectores adeno-asociados recombinantes (rAAV), ampliamente extendidos para aplicaciones en TG. Los virus adenoasociados (AAV) son virus defectivos de la familia *Parvoviridae*, con cápside protéica icosaédrica. Su genoma consiste en una molécula lineal de DNA monocatenario¹⁴⁹, de aproximadamente 4.7 Kb, flanqueada en sus extremos por repeticiones terminales invertidas (ITRs)¹⁵⁰, que sirven como origen de replicación viral y como señal de empaquetamiento¹⁵¹, y que están compuestas por 145 nucleótidos

(nt), de los cuales los primeros 125 nt autoaparean, formando una horquilla en forma de T¹⁵².

Los AAV infectan de forma eficiente tanto células en división, como células quiescentes, y son capaces de integrarse en loci específicos del genoma (sitios de integración de AAV), permaneciendo en latencia en la célula infectada en ausencia de un virus auxiliar (un adenovirus o un herpesvirus) que permita su replicación¹⁵³.

En los rAAV, las secuencias de origen viral son reemplazadas por el transgén terapéutico, a excepción de los ITRs, que se mantienen en para su replicación y empaquetamiento durante el proceso de producción de vectores. La eliminación de los genes virales hace que los vectores derivados de los AAV no se integren en el genoma y, en caso de hacerlo, esto ocurre con muy baja frecuencia y de forma inespecífica.

Interacciones entre la cápside de AAV y receptores de la superficie celular determinan su internalización y, en consecuencia, su infectividad. Hasta la fecha, se han identificado al menos 12 serotipos de AAV humanos, con diferente tropismo. Como sucede en el caso de AdVs, un alto porcentaje de la población presenta anticuerpos neutralizantes que reconocen distintos epítomos de las cápsides de AAVs, debido a la exposición natural a ellos. Con el fin de evadir una respuesta inmunitaria humoral, e incrementar la infectividad y el tropismo de los rAAV, se han creado librerías de cápsides quiméricas, combinando propiedades de distintos serotipos mediante diseño racional y evolución dirigida^{153,154}

La limitada capacidad de carga de los rAAV, y que la expresión del transgén dependa de la conversión previa del genoma viral en una molécula de DNA bicatenario, son otros de los inconvenientes de los rAAV. En AAV no modificados (cepa salvaje), la

conversión del genoma viral en una molécula de dsDNA se produce, tanto mediante síntesis *de novo*, como mediante hibridación de genomas complementarios procedentes de AAV independientes que han infectado la misma célula¹⁵³. Para solventar estas limitaciones, se han desarrollado estrategias como la generación de rAAV duales para vehicular cassettes de expresión largos, o rAAV autocomplementarios.

1.5.1.3 Vectores retrovirales (RVs)

Los virus que pertenecen a la familia *Retroviridae*, entre los que se encuentran gammaretrovirus y lentivirus, se caracterizan por contener dos moléculas de RNA de cadena sencilla y polaridad positiva que, tras sufrir un proceso de retrotranscripción, se incorporan en el genoma de la célula hospedadora.

Los vectores retrovirales (RVs) fueron los primeros en ser aplicados en TG para la corrección de desordenes genéticos humanos. En los primeros ensayos clínicos se usaron vectores gammaretrovirales (g-RVs) basados en el virus de la leucemia murina (MLV) para tratar el Síndrome de inmunodeficiencia combinada severa por déficit de adenosina desaminasa (ADA-SCID), mediante la transducción *ex vivo* y posterior reinfusión de células T¹⁵⁵ y HSCs¹⁵⁶⁻¹⁵⁸. Sin embargo, en estos primeros pacientes no se logró corregir la inmunodeficiencia, y fue necesario mantener la terapia de reemplazo enzimático debido, al menos en parte, a la baja eficiencia de transducción¹⁵⁹.

La implementación de notables mejoras en los protocolos de transducción permitió poner en marcha, entre 1999 y 2009, cinco ensayos clínicos en los que se trataron con g-RVs un total de 20 pacientes con inmunodeficiencia combinada severa ligada al X (SCID-X1), debida a mutaciones en el gen IL2RG, y 27 pacientes con ADA-SCID¹⁵⁹⁻¹⁶⁴. La inmunodeficiencia se consiguió corregir sustancialmente en 17 de los pacientes

SCID-X1, y 19 de los ADA-SCID, demostrando el potencial terapéutico de la corrección de células autólogas. Desafortunadamente, estos ensayos también revelaron la genotoxicidad de los g-RVs, al provocar el desarrollo de leucemias en 5 de los 20 pacientes SCID-X1, hecho que posteriormente se relacionó con actividad enhancer en las repeticiones terminales largas (LTR) del vector, y con una tendencia de g-RVs a integrarse en genes transcripcionalmente activos, como proto-oncogenes¹⁵⁹.

1.5.1.3.1 Vectores lentivirales (LVs)

Los lentivirus son parte de la familia de retrovirus que incluye el virus de inmunodeficiencia humana 1 (VIH-1) y se caracterizan por un largo período de incubación (Lente significa lento en latín). Como otros retrovirus, los lentivirus infecciosos contienen en su genoma los 3 genes principales (gag-pol-env) que codifican proteínas estructurales y funcionales. Además, el genoma del RNA lentiviral codifica varios genes accesorios (dependiendo del virus; vif, vpr, vpu, nef, etc.) involucrados en la regulación transcripcional y la infectividad. Los lentivirus son los únicos retrovirus que infectan de manera eficiente las células que no se dividen. Esta propiedad los ha convertido en una de las mejores plataformas para desarrollar vectores de TG.

Los vectores lentivirales (LVs, del inglés lentiviral vectors) ofrecen varias ventajas con respecto a otros sistemas por el hecho de ser muy eficientes para la modificación genética de células primarias, incluidas neuronas, HSCs y células T. De todos los LVs, los basados en VIH-1 son los más utilizados para las estrategias de TG clásica. La última generación de LVs está desprovista de todas las secuencias codificantes de VIH. Las proteínas necesarias para la formación de partículas y las actividades enzimáticas se suministran en trans (mediante plásmidos separados que codifican las proteínas Gag, Pol, Env, Tat y Rev). El transactivador Tat también es prescindible para la generación de

partículas LV en las últimas generaciones de LVs, donde el promotor U3 del LTR en 5' se reemplaza por un promotor fuerte (p.e. CMVp), lo que da como resultado una transcripción independiente de Tat.

Los LVs son los vectores más eficaces para la modificación génica de las HSCs. En la actualidad, hay mas de 278 ensayos clínicos aprobados que utilizan vectores lentivirales (<http://www.abedia.com/wiley/vectors.php>) para el tratamiento del cáncer (175), enfermedades monogénicas (70), enfermedades infecciosas (20) y otras enfermedades. Estos estudios han resultado, por ahora, en la aprobación como medicamento de TG a Kymriah, para el tratamiento de niños con leucemia linfoblástica aguda (ALL).

A pesar de estos logros, los LVs siguen insertándose al azar dentro de regiones transcripcionalmente activas, por lo que existe un riesgo inherente a la modificación genética cuando se utilizan dichos vectores. Además, no son útiles para aplicaciones terapéuticas utilizando herramientas de EG, dado que la expresión continuada de las NE no es recomendable.

1.5.1.3.2 Vectores lentivirales no integrativos (IDLVs)

El uso de vectores lentivirales no integrativos (IDLVs, del inglés Integration-defective lentiviral vectors) es una opción lógica para minimizar el riesgo de mutagénesis insercional cuando las células diana son células quiescentes. Los IDLVs son, también, una atractiva alternativa para proporcionar una expresión transitoria de transgenes en células en división, incluyendo los sistemas de EG como ZFNs y CRISPR/Cas. Al igual que en el caso de LVs, el tropismo de IDLVs se puede modificar, y adaptar en función de las células diana, pseudotipando con diferentes proteínas para la envuelta.

La principal desventaja de IDLVs respecto a LVs es que alcanzan menores niveles de expresión génica. Esta reducción en los niveles de expresión se debe, fundamentalmente, a que los IDLVs sufren silenciamiento epigenético como resultado de cromatinización nuclear. En 2013, el Pelascini et al. observaron que la desacetilación de histonas constituía el principal mecanismo relacionado con el silenciamiento de IDLVs.

Para mejorar la expresión de sistemas de transferencia génica episomales, tanto virales como no virales, se han empleado diferentes estrategias. En sistemas de transferencia génica episomales no virales con frecuencia se ha recurrido al uso de secuencias de unión a la matriz nuclear (SARs o S/MARs) para mejorar los niveles de expresión y mantener esta expresión a lo largo del tiempo. Esto es debido, fundamentalmente, a la capacidad de los elementos SARs de unirse a factores como SATB1 (proteína 1 de unión a secuencias ricas en AT), Nmp4 (proteína de unión a matriz nuclear), y CTCF (factor de unión a CCCT). Además, por su capacidad de agregar a las nucleoproteínas, reclutar a las histona acetiltransferasas (HATs) y a los complejos de remodelación de la cromatina dependiente de ATP¹⁶⁵. Al igual que los SARs, la inclusión de los elementos derivados de HS4 (5' cHS4, del inglés chicken hypersensitive site-4) han demostrado su eficacia en incrementar la expresión en sistemas episomales¹⁶⁶, debido en parte a la interacción de los elementos cHS4 con la matriz nuclear vía proteínas nucleares CTCF¹⁶⁷. Estos factores nucleares evitan la heterocromatinización de estos elementos episomales dentro de los núcleos de las células tratadas¹⁶⁸⁻¹⁷¹. En nuestro laboratorio, hay una línea de investigación muy activa en este sentido. Hemos demostrado previamente que la inclusión de elementos (IS2), los cuales contienen un elemento sintético SAR con un fragmento de cHS4, mejora el patrón de expresión de los LVs integrativos. Sin embargo, este efecto mostraba una gran dependencia del tipo celular usado en cada experimento¹⁷².

1.6 SÍNDROME DE WISKOTT-ALDRICH (WAS)

El Síndrome de Wiskott-Aldrich (WAS) es una inmunodeficiencia primaria, que afecta casi exclusivamente a varones. Los afectados muestran inmunodeficiencia celular y humoral, infecciones recurrentes, eczema, microtrombocitopenia con tendencia al sangrado, y propensión al desarrollo de enfermedades autoinmunes y leucemias^{173,174}. Su incidencia se estima en 1-4:10⁶ y, en ausencia de tratamiento, la esperanza de vida se sitúa en torno a los 15 años¹⁷⁵. Fue en 1937 cuando el Dr. Wiskott refería, por primera vez, el caso de tres hermanos que cursaban con los síntomas característicos de esta enfermedad. Más tarde, en 1954, los Drs. Aldrich, Steinberg y Campbell establecían, mediante el estudio de 9 generaciones de una familia de ascendencia holandesa, con varios varones afectados, que este síndrome tiene bases genéticas, con un patrón de herencia recesivo ligado al cromosoma X¹⁷⁶. Finalmente, en 1994 Derry et al. hallaron, mediante estudios de ligamiento, el gen cuyas mutaciones causan tres fenotipos distintos: WAS, trombocitopenia ligada al cromosoma X (XLT) y, con menor frecuencia, neutropenia ligada al X (XLN)¹⁷⁷⁻¹⁸¹. Este gen, denominado gen *WAS*, se localiza en el brazo corto del cromosoma X, en posición Xp11.22-p11.23, y se compone por 12 exones y 1823 pares de bases (pb)^{177,182}.

El gen *WAS* codifica la proteína WAS (WASP), de expresión específica y constitutiva del linaje hematopoyético, y con un papel clave en la señalización y reorganización del citoesqueleto de actina en células hematopoyéticas¹⁸³⁻¹⁸⁵. Hasta el momento se han descrito más de 400 mutaciones para *WAS*, registradas en la base de datos de mutaciones génicas humanas (<http://www.hgmd.cf.ac.uk>), la mayoría de las cuales son únicas. Esto significa que, por lo general, cada familia afectada tiene su propia mutación característica. Predominan mutaciones puntuales en los primeros 4 exones del gen, mutaciones que afectan a sitios de splicing en los intrones 6 a 11, e inserciones y deleciones que resultan en un final de la traducción prematuro. También se han

descrito mutaciones complejas, que incluyen dobles mutaciones puntuales, o combinaciones de inserciones y deleciones, aunque estas son más raras¹⁷³. En mujeres portadoras de una mutación para *WAS* en heterocigosis, hay una selección preferencial por el cromosoma X salvaje en células hematopoyéticas, de forma que el cromosoma X que porta la mutación se inactiva no aleatoriamente, salvo en casos esporádicos¹⁸⁶⁻¹⁸⁹.

Las mutaciones en *WAS* provocan, generalmente, ausencia o niveles reducidos de la proteína. En algunos pacientes en que la proteína está presente en niveles reducidos, WASP está truncada y no es funcional. En cuanto a la gravedad de las manifestaciones, depende de la mutación en el gen *WAS*, asociándose los fenotipos más graves con mutaciones que provocan la ausencia de proteína¹⁷³.

WASP es una proteína de 502 aminoácidos, multimodular, perteneciente a la familia de promotores de nucleación de la actina¹⁹⁰, a la que pertenecen también las proteínas WASP neuronal (N-WASP) y SCAR/WAVE^{191,192}. Las proteínas de esta familia están implicadas en numerosos procesos, incluyendo, entre otros, la transducción de señales extracelulares relacionadas con la movilización y reorganización del citoesqueleto de actina^{183,185,190,192-194}. WASP tiene un dominio regulador y otro efector, que le permiten participar en la polimerización de filamentos de actina dependiente de señales extracelulares. La ausencia de WASP provoca la desregulación de la polimerización de los filamentos de actina, directamente relacionada con las manifestaciones fenotípicas de WAS. El citoesqueleto de actina juega un papel importante en muchas respuestas celulares, entre ellas, relacionadas con endocitosis, tráfico vesicular, migración, reorganización de la membrana para la señalización celular, y sinapsis inmunológica. La carencia de WASP perturba la localización, activación y función de los leucocitos, creando respuestas inmunitarias adaptativas e innatas defectuosas¹⁸⁵.

La expresión de WASP confiere ventaja selectiva en determinados linajes hematopoyéticos, especialmente en células T reguladoras (Treg) y en células T natural killer (NKT)¹⁹⁵.

1.6.1 Terapias para el tratamiento de WAS

La terapia actual para el síndrome de Wiskott-Aldrich incluye el tratamiento de las infecciones, la autoinmunidad, y la trombocitopenia. El tratamiento de las infecciones se lleva a cabo mediante la administración de antibióticos y de inmunoglobulinas (Igs)^{182,196}. De esta manera, los pacientes con infecciones recurrentes, bajos niveles de Igs, o respuestas inmunológicas atenuadas, son tratados con dosis de reemplazo de Igs intravenosas (IVIG), para restablecer los niveles fisiológicos de las Igs. En cuanto a las manifestaciones de autoinmunidad, los tratamientos inmunomoduladores, incluyendo IVIG, mejoran levemente los síntomas. Agentes inmunosupresores, como corticosteroides, son ampliamente utilizados. Sin embargo, su toxicidad (incremento en el riesgo de infecciones oportunistas) limita su uso generalizado^{182,197,198}. Además, aún se tienen pocos datos sobre el uso de otros agentes inmunomoduladores como ciclosporina, azatioprina, o ciclofosfamida. La trombocitopenia, asociada a la enfermedad, es tratada mediante la transfusión de plaquetas con el fin de limitar el riesgo de hemorragias. En algunos casos, para evitar este riesgo se procede a realizar una esplenectomía, proceso que mejora y restablece el número de plaquetas. Sin embargo, se ha relacionado la esplenectomía con un incremento de la mortalidad en pacientes que posteriormente se someten a un trasplante de células troncales hematopoyéticas (HSCs)^{199,200}. El eczema, relacionado con la enfermedad, puede ser tratado de forma tópica, con corticosteroides. Estos tratamientos se dirigen a paliar la sintomatología de WAS, pero no posibilitan el restablecimiento del sistema hematopoyético e inmunitario.

Hasta la fecha, el trasplante alogénico de células madre hematopoyéticas CD34⁺ obtenidas de médula ósea es el único tratamiento curativo para WAS.

El primer trasplante de médula ósea para el tratamiento de WAS, se llevó a cabo en 1978, utilizando un donante haploidéntico²⁰¹. El donante ideal es un familiar con HLA compatible al del paciente, reduciéndose el éxito del tratamiento cuando el donante es un familiar con HLA no compatible, o un donante con HLA compatible, pero no emparentado. A pesar de que actualmente la supervivencia es muy elevada en todos los tipos de trasplante (en torno al 90%), permanece el riesgo de que surjan complicaciones. Este riesgo varía en función del tipo de donante, y es menor en trasplantes de hermanos histocompatibles. El éxito del trasplante también se ha asociado a la edad del paciente, de forma que incrementa si se realiza dentro de los 5 primeros años de vida^{197,202}. No obstante, la falta de donantes y los efectos secundarios en muchos pacientes han contribuido a la búsqueda de terapias alternativas.

La TG mediante el trasplante con HSCs expresando el gen *WAS* ha sido la estrategia común utilizada. Esta estrategia ha dado lugar a los primeros beneficios clínicos en pacientes con WAS^{203,204}. El primer ensayo clínico de TG de WAS fue realizado por el grupo del Dr Klein en Alemania utilizando γ -RVs con resultados positivos en cuanto a restauración de los defectos inmunológicos y de la trombocitopenia²⁰³, pero con un desarrollo de leucemias en la mayoría de los pacientes (7/10)²⁰⁵. El segundo ensayo clínico publicado fue realizado por el grupo del Dr Naldini en Italia y utilizó vectores lentivirales. En este ensayo no se ha observado ninguna evidencia de transformación tumoral y los resultados en cuanto eficiencia terapéutica han sido muy satisfactorios, salvo problemas en conseguir una restauración completa de la microtrombocitopenia²⁰⁴. Otros ensayos clínicos para TG de WAS utilizando vectores lentivirales similares a los del Dr Naldini (Grupo internacional Francia-Inglaterra-EEU) confirman la baja genotoxicidad de estos vectores, la eficiencia terapéutica en la mayor parte de los defectos funcionales y los problemas para conseguir restaurar la

microtrombocitopenia. Nuestro grupo ha desarrollado vectores optimizados que están siendo propuestos en la actualidad como medicamento huérfano para el tratamiento de WASP²⁰⁶⁻²⁰⁸.

Sin embargo, a pesar del potencial de los vectores lentivirales como herramienta terapéutica (altamente eficiente y poco genotóxicos), su inserción al azar en el genoma es potencialmente deletérea para las HSCs. De hecho, estudios recientes del grupo del Dr Montini han demostrado que, incluso los vectores lentivirales con un mejor perfil de seguridad son potencialmente tumorigénicos en determinadas condiciones²⁰⁹. Es por tanto necesario investigar alternativas a los vectores lentivirales que puedan ofrecer eficiencias similares pero que eviten los potenciales efectos de genotoxicidad. Tal y como se ha descrito previamente, una alternativa que está emergiendo en los últimos años se basa en la edición génica mediante nucleasas específicas.

JUSTIFICACIÓN

Los grandes avances que están teniendo lugar en el campo de la EG, particularmente desde la irrupción del sistema CRISPR en este, están cambiando la forma de abordar la investigación en biología molecular y biomedicina, y plantean un escenario donde estrategias basadas en EG se presenten como alternativa terapéutica para el tratamiento de determinadas enfermedades. Principalmente, podrían verse beneficiadas por este tipo de estrategias ciertas enfermedades monogénicas que hasta la fecha solo cuentan con tratamientos dirigidos a paliar su sintomatología, o para las que acceder a un tratamiento curativo no siempre es posible. En este tipo de afecciones, la reparación del gen alterado mediante EG representaría una alternativa prometedora para revertir la enfermedad.

Una de las enfermedades para cuyo tratamiento podrían diseñarse estrategias de EG en un futuro es el Síndrome de Wiskott-Aldrich (WAS), inmunodeficiencia primaria de herencia recesiva ligada al cromosoma X, causada por mutaciones en el gen *WAS*, y asociada a microtrombocitopenia, eczema, infecciones recurrentes y autoinmunidad. Actualmente el único tratamiento curativo para pacientes con WAS es el trasplante alogénico de HSCs de médula ósea. No obstante, la falta de donantes compatibles para algunos pacientes ha contribuido a la búsqueda de terapias alternativas. Para este tipo de pacientes existe la posibilidad de infusión de células modificadas genéticamente *ex vivo*. El uso de LVs capaces de expresar el gen correcto en HSCs es la principal estrategia mediante la que se consigue esta modificación *ex vivo*. Sin embargo, el desarrollo de estrategias que permitan una expresión más fisiológica de la proteína terapéutica sigue siendo apremiante. La estrategia ideal sería, en este sentido, corregir mediante EG la mutación causante de WAS.

Un punto fundamental para que una estrategia de EG sea trasladable a clínica es conseguir un balance entre eficiencia y especificidad que permita la modificación de las células diana de una forma eficiente, sin supeditar a esta eficiencia la especificidad.

Es decir, ante todo, la estrategia debe garantizar la seguridad. En relación con este aspecto, uno de los principales problemas de este tipo de abordaje es que las NE pueden generar roturas, no solo en la diana, sino en otras localizaciones del genoma con cierto grado de homología con respecto a la secuencia diana. Además, en aquellas estrategias que requieren de un DNA donador para su inserción mediante RH, también se ha descrito que este donador puede llegar a insertarse al azar en el genoma, en sitios en los que se hayan generado roturas (tanto espontáneas, como generadas de forma inespecífica por las NE utilizadas). Por ello, es imprescindible evaluar en términos de eficiencia y seguridad todos los componentes que se requieren para la EG, así como la forma de hacer llegar estos componentes a las células diana. Esto incluye la evaluación, tanto del sistema de entrega de la NE a las células diana, el cual ha de proporcionar una expresión transitoria de la NE (Cas9, en los sistemas CRISPR tipo II), como del DNA donador, en aquellas estrategias que lo requieran

HIPÓTESIS

Nuestra hipótesis es que la mejora de la eficiencia y la seguridad de las herramientas de EG permitirán, en un futuro cercano, la corrección génica en trastornos con base genética, como WAS.

OBJETIVOS

1. Mejorar la entrega de los componentes de los sistemas CRISPR/Cas9 a las células diana.
 - a. Evaluar la eficiencia de distintos métodos de entrega del sistema CRISPR/Cas9 en modelos celulares: K562.
 - b. Evaluar, en términos de eficiencia, distintos métodos de entrega del sistema CRISPR/Cas9 en HSCs y en células T.
 - c. Desarrollo de nuevos sistemas de entrega más eficaces basados en IDLVs.

2. Desarrollo de modelos celulares para estudiar eficiencia y seguridad de la reparación génica mediante el sistema CRISPR/Cas9.
 - a. Generación de líneas hematopoyéticas reporteras, que permitan comparar eficiencia y especificidad de RH.
 - b. Diseño de donadores que permitan analizar eficacia y seguridad de RH en los modelos generados.
 - c. Evaluar eficiencia de RH frente a inserciones inespecíficas del donador, utilizando distintos sistemas de entrega de las NE.
 - d. Evaluar eficiencia de RH frente a inserciones inespecíficas del donador, utilizando distintos donadores.

3. Análisis de eficiencia y seguridad de los sistemas CRISPR/Cas9 específicos para *WAS* en HSCs y células T.
 - a. Eficacia y seguridad de corte en el locus *WAS*
 - b. Eficacia y seguridad de RH en el locus *WAS*.

MATERIAL Y MÉTODOS

1.7 TIPOS CELULARES Y CONDICIONES DE CULTIVO

1.7.1 Línea celular HEK 293T

Las células HEK 293T (ATCC® CRL-11268™) son células de morfología epitelial, procedentes de riñón de feto humano, que expresan el antígeno T del virus SV40, y son utilizadas, por su alta transfectabilidad, como células empaquetadoras en la producción de LVs e IDLVs. Se cultivan en medio DMEM con alta concentración de glucosa (DMEM High Glucose, Biowest), suplementado con un 10% de suero fetal bovino (FBS)(Biowest), a 37°C y 10% de CO₂.

1.7.2 Línea celular K562

(ATCC® CCL-243™). Línea celular eritroleucémica derivada de una paciente con leucemia mieloide crónica (LMC), que presenta un cariotipo triploide, con disomía para el cromosoma X. Estas células se cultivaron en RPMI 1640 (Biowest) suplementado con un 10% de FBS (Biowest), y Penicilina/Estreptomina (P/S) (Biowest) a una concentración de 10ml/l, a 37°C, en una atmósfera del 5% de CO₂.

1.7.3 Línea celular Jurkat

Línea celular derivada de leucemia linfoblástica aguda de células T, mantenida en RPMI 1640 (Biowest), suplementado con un 10% de FBS (Biowest), y P/S (Biowest) a una concentración de 10ml/l, a 37°C y un 5% de CO₂.

1.7.4 Células progenitoras neuronales humanas (NPCs)

Células multipotentes derivadas de células madre embrionarias humanas (hESCs) H9 (WA09), capaces de diferenciarse en neuronas, oligodendrocitos, y astrocitos. Estas células se mantuvieron a 37°C, en una atmósfera del 5% de CO₂, en medio sin suero

StemPro NSC SFM (Gibco), conteniendo los factores de crecimiento FGF y EGF, a una concentración de 5µg/l en ambos casos. Para favorecer la adhesión de las NPCs, las placas de cultivo se revistieron previamente con los sustratos Poli-L-Ornitina (Sigma-Aldrich, Ref. P4957) a una concentración de 10 µg/ml, y Laminina (ThermoFisher Scientific, Ref. 23017-015) a una concentración de 20 µg/ml.

1.7.4.1 Diferenciación neuronal *in vitro* y cultivo de neuronas diferenciadas

La diferenciación neuronal *in vitro* se llevó a cabo retirando del medio de cultivo de las NPCs los factores de crecimiento FGF y EGF, una vez estas habían alcanzado una confluencia del 90%. Las células en diferenciación se mantuvieron en cultivo por un periodo de 28 días, durante los cuales se realizaron cambios de medio cada 3-4 días.

1.7.5 Células troncales pluripotentes inducidas (iPSCs)

La línea de iPSCs PBMC1-iPS4F1, generada y caracterizada por Montes et al.²¹⁰, se cultivó sobre células mesenquimales estromales (MSCs) irradiadas (Inbiobank), en medio PSC (pluripotent stem cell medium), consistente en KnockOut DMEM (Life Technologies) suplementado con un 20% de suero de reemplazo (KnockOut serum replacement, Life Technologies), 8ng/ml de factor básico de crecimiento de fibroblastos (bFGF2, Milteny Biotec), inhibidor de ROCK Y-27632 (Sigma-Aldrich) a una concentración de 10 µM, glutamina 1mM, 1% de aminoácidos no esenciales, y 0.1mM de β-mercaptoetanol.

1.7.6 Cultivos celulares primarios de mucosa oral humana y fibroblastos de piel

Los cultivos de células primarias de mucosa oral humana (HOM) y fibroblastos de piel (HSF) se obtuvieron a partir de biopsias de donantes sanos. Las biopsias de piel y mucosa oral se lavaron en PBS 1x y se digirieron enzimáticamente con 2 mg/ml de

Colagenasa I de *Clostridium histolyticum* (Gibco), a 37 °C durante 6 horas. Los fibroblastos aislados se recogieron por centrifugación y se expandieron en medio de cultivo basal (DMEM complementado con un 10% de FBS, y 100 U/ml de penicilina, 0,1 mg/ml de estreptomycin y 0,25 µg/ml de anfotericina B, todas de Sigma-Aldrich, bajo condiciones estándar de cultivo celular).

1.7.7 Aislamiento de células troncales hematopoyéticas CD34+ (HSCs)

El aislamiento de HSCs se realizó a partir de sangre periférica movilizada. Las muestras se obtuvieron a través del Servicio de Hematología del Hospital Reina Sofía de Córdoba, tras obtener el consentimiento informado de los donantes. Tras la recepción de las muestras, se realizaron dos lavados para la eliminación de plaquetas (diluyendo la sangre 1:4 en PBS, y centrifugando a 300g durante 10min, con freno). Entre ambos lavados, las células se resuspendieron en 2ml de Quicklysis erythrocyte lysing solution (Cytognos), y se incubaron durante 10min a RT, para la lisis de eritrocitos. Transcurrido el tiempo de incubación, para retirar la solución de lisis se centrifugaron las células a 300g durante 10min, con freno. Posteriormente, las HSCs fueron aisladas mediante separación magnética (AutoMACS, Miltenyi Biotec), usando el CD34 MicroBead Kit UltraPure (Miltenyi Biotec) para selección positiva de células expresando el antígeno humano CD34, y siguiendo las recomendaciones del fabricante. Una vez aisladas, las células enriquecidas se tiñeron con anticuerpo anti-CD34 humano conjugado con Ficoeritrina-Cianina 7 (anti-CD34-PECy7, eBiosciencie). Finalmente, mediante citometría de flujo (FACSCanto II, Becton Dickinson), se evaluó la pureza, que en todos los casos fue superior al 95%.

1.7.7.1 Cultivo de HSCs (CD34+)

Las HSCs se cultivaron a 37°C en una atmósfera del 5% de CO₂, en StemSpan SFEM (Stemcell Technologies), suplementado con las citoquinas Stem Cell Factor (SCF) (100

ng/ml), Flt3-Ligando (100 ng/ml), Trombopoyetina (TPO) (100 ng/ml), Interleucina-6 (IL-6) (100 ng/ml), StemRegenin 1 (SR1) (0,75 μ M), Prostagladina E2 (PGE2) (10 μ M), de la casa comercial Peprotech, y UM171 (Stemcell Technologies) (35 nM), para el mantenimiento del estado indiferenciado.

1.7.8 Linfocitos T primarios aislados a partir de sangre periférica movilizada

El aislamiento de células T humanas se realizó a partir de muestras de sangre periférica movilizada procedentes del Servicio de Hematología del Hospital Reina Sofía de Córdoba, tras obtener el consentimiento informado de donantes sanos. Tras el aislamiento de HSCs CD34+, descrito anteriormente, a partir de la fracción negativa obtenida se aislaron células T mediante separación magnética negativa (MidiMACS separator, Miltenyi Biotec), usando el cóctel de anticuerpos frente a las células no diana del Pan T Cell isolation kit (Miltenyi Biotec), y siguiendo las instrucciones de la casa comercial. Las células T se cultivaron a 37°C, 5% CO₂, en medio TexMACS (Miltenyi Biotec) suplementado con un 5% de suero humano (Biowest), y 100 IU/ml de Interleucina-2 recombinante humana (IL-2) (Peprotech). Las células T fueron estimuladas con anti-CD3 y anti-CD28 (T Cells Trans Act, Miltenyi Biotec), utilizando los volúmenes recomendados por la casa comercial en base a la densidad celular, 48h antes de ser nucleofectadas o transducidas.

1.8 SISTEMAS UTILIZADOS PARA ENTREGA DE NE Y DNA DONADORES

La entrega del sistema CRISPR/Cas9 a las células diana (K562, HSCs, y células T) se realizó por tres métodos distintos: en forma de plásmido o de IDLV, con las construcciones Lgl₁Cas9 o Lgl₉Cas9 como base, o en forma de ribonucleoproteína ensamblada *in vitro* previamente. Los DNA donadores se entregaron como plásmido,

o como producto de PCR, mediante nucleofección. Los distintos sistemas empleados se describen a continuación.

1.8.1 Producción y titulación de LV e IDLV de segunda generación

La producción de LV e IDLV se llevó a cabo mediante la co-transfección de células HEK 293T con tres plásmidos: un plásmido de empaquetamiento, conteniendo los genes virales Gag, Pol, Rev y Tat, un plásmido codificando las proteínas de la envuelta, y el correspondiente plásmido vector. El plásmido de empaquetamiento que se utilizó es el pCMVR8.74 (http://www.addgene.org/Didier_Trono), en el caso de LV, y el pCMVDRD8.74 en el caso de IDLV (cedido generosamente por el doctor Steven J Howe, desde el University College London - Institute of Child Health). Para la envuelta se empleó el plásmido pMD2.G (VSV-G) (http://www.addgene.org/Didier_Trono), codificando la proteína G del virus de la estomatitis vesicular (VSV-G), cuyo uso para el pseudotipado de LV está ampliamente extendido²¹¹.

Las células empaquetadoras HEK 293T se plaquearon 16h antes de la transfección en placas Petri de 10cm de diámetro, recubiertas con amina (PureCoat Amine Dishes, Corning) para favorecer la adhesión celular y la proliferación. Para conseguir una confluencia del 80-90% en el momento de la co-transfección, se sembraron 5×10^6 células por placa.

Los tres plásmidos con los que se co-transfectaron las 293T se resuspendieron en medio DMEM sin suero, y se mezclaron con LipoD293™ (SignaGen Laboratories), llevando a cabo la co-transfección tal y como ha sido previamente descrito²¹². Los sobrenadantes conteniendo las partículas virales se recogieron y filtraron transcurridas 48h tras la transfección, y se concentraron mediante ultracentrifugación (Sorvall VX Ultra, Thermofisher), centrifugando a 23000rpm y 4°C durante 2h.

La siguiente tabla muestra la cantidad empleada de cada uno de los plásmidos para la producción de LV e IDLV respectivamente:

Tabla 3. Cantidad empleada de cada uno de los plásmidos requeridos en la producción de LVs e IDLVs de segunda generación (1- plásmido de empaquetamiento, 2- plásmido pMD2.G (VSV-G) para la envuelta, 3- plásmido vector, codificando el transgén), y del agente de transfección (en este caso, LipoD293).

Tipo de Vector	Plásmido	μg	μl LipoD293™
LV	Plásmido vector	9	45
	pCMVR8.74	7	
	pMD2.G (VSV-G)	3	
IDLV	Plásmido vector	14	60
	pCMVDRD8.74	7	
	pMD2.G (VSV-G)	3	

Los títulos virales, entendidos como la estimación de las unidades de transducción por ml (UT/ml), se determinaron mediante RT-qPCR empleando el kit para titulación de vectores lentivirales de la casa comercial abm (abm® Ref. LV900), que permite estimar este valor a partir del número de genomas virales por ml (GC/ml).

1.8.1.1 Transducción de diferentes tipos celulares con LV o IDLV

Para la transducción de los distintos tipos celulares con los sobrenadantes virales de LV o IDLV, las células fueron previamente lavadas con PBS (Duldecoco's PBS 1x, biowest), contadas, y plaqueadas en función del número de células y tamaño celular. En el caso de células adherentes (HEK 293T y NPCs), para su desprendimiento las células fueron disociadas con TrypLE 0,4x (Gibco, Thermo Fisher Scientific) tras el lavado con PBS. Las transducciones se llevaron a cabo incubando las células durante 5h con distinta cantidad de partículas virales en función de la multiplicidad de infección (MOI)

deseada. El MOI hace referencia al ratio de partículas virales por célula, y se estimó en cada caso en base al título determinado mediante RT-qPCR.

La entrega del sistema CRISPR/Cas9 en forma de IDLV se llevó a cabo mediante la transducción de las células diana (K562, HSCs, y células T estimuladas) con los sobrenadantes virales, usando distintos MOI en base a las UT/ml estimadas mediante RT-qPCR.

1.8.2 Formación de ribonucleoproteína Cas9/gRNA *in vitro*

El complejo RNP se formó incubando 120 pmol de gRNA sintético, incorporando en ambos extremos la modificación 2'-O-metil-3' fosforotioato (Synthego)²¹³, con 104 pmol de la proteína Cas9 (Integrated DNA Technologies), en un volumen final de 5 μ l, durante 20' a 37°C. Una vez formado el complejo RNP, este se entregó a las células diana mediante nucleofección.

1.8.3 Nucleofección

La transferencia del sistema CRISPR/Cas9 en forma plásmido, así como de RNP, se llevó a cabo por nucleofección, tanto en nuestras células diana finales (HSCs y células T), como en las células modelo (K562). En el caso de K562, se empleó el kit de nucleofección SF Cell Line (Lonza, Referencia: V4XC-2032) para la nucleofección de 2×10^5 células en el 4D-Nucleofector de Amaxa (Lonza), usando el programa FF-120. El kit de nucleofección P3 Primary Cell (Lonza, Referencia: V4XP-3032) se usó para la nucleofección de 10^5 células CD34⁺ (HSCs) en el 4D-Nucleofector de Amaxa (Lonza), con el programa E0-100, así como para la nucleofección, con el programa E0-115, de 8×10^5 células T, estimuladas con anti-CD3 y anti-CD28 (T Cells Trans Act, Miltenyi Biotec) 48h antes de la nucleofección.

1.9 GENERACIÓN DE MODELOS CELULARES PARA ESTUDIAR EFICACIA Y SEGURIDAD DE LA EG

1.9.1 Generación de la línea celular reportera K562 SE- δ GF-I₁-P

La línea celular K562 SE- δ GF-I₁-P, conteniendo en su genoma múltiples integraciones de la secuencia codificante de eGFP, mutada e interrumpida por 387pb del primer intrón del gen *WAS*, bajo el promotor SFFV, se generó mediante la transducción de K562wt con un MOI aproximado de 20 del LV SE- δ GF-I₁-P.

Para ello, previamente se diseñó el vector SE- δ GF-I₁-P, cuya obtención implicó varios subclonajes. Por un lado, con la pareja de primers hWASP2 Fw / hWASP2 Rv, se amplificó mediante PCR a partir de DNA genómico de células K562wt, un fragmento de 811pb del gen *WAS*, que incluía las dianas para Cas9 localizadas en el primer intrón del gen. Este fragmento de PCR se clonó en el plásmido pCR2.1-TOPO (pCR2.1 TA Cloning Kit, Thermofisher), obteniendo así el plásmido pCR2.1 WAS811. A continuación, a partir del plásmido pCR2.1 WAS811 se amplificó un fragmento del primer intrón de *WAS* de 387pb, con la pareja de primers hWASP-I₁Pfo Fw / hWASP-I₁Pfo Rv, aportando la diana de la enzima de restricción *PfoI* a ambos lados del fragmento. Una vez más, el producto de PCR se clonó en el plásmido pCR2.1-TOPO, originando así el plásmido pCR2.1 WAS I₁ 387.

Por otro lado, mediante mutagénesis dirigida por PCR se eliminó el primer ATG de la secuencia codificante de eGFP, utilizando la pareja de primers *BamHI*- δ GFP Fw y *XhoI*-GFP Rv, y el plásmido SE LV como molde. El primer *BamHI*- δ GFP Fw se diseñó de forma que hibridase aguas abajo del primer ATG. A su vez, estos primers aportaban al fragmento de amplificación las dianas de restricción *BamHI*, en el extremo 5', y *XhoI* en el extremo 3', facilitando su manipulación posterior. De nuevo, el producto de PCR

se clonó en el plásmido pCR2.1-TOPO, generando así el plásmido pCR2.1 δ GFP. El fragmento de 387pb del primer intrón de *WAS* contenido en el plásmido pCR2.1 WAS I₁ 387 se clonó en pCR2.1 δ GFP, tras digerir ambos plásmidos con la enzima de restricción *PfoI*, obteniendo así el plásmido pCR2.1 δ GF-WAS I₁-P. Finalmente, la inclusión del cassette SFFV_ δ GF- WAS I₁-P en un entorno lentiviral se llevó a cabo de forma dirigida, mediante digestión doble con las enzimas *BamHI* y *XhoI* de los plásmidos pCR2.1 δ GF-WAS I₁-P y SE LV, conteniendo este último el cassette de expresión SFFV-eGFP original, y posterior ligación del fragmento δ GF-WAS I₁-P en el esqueleto lentiviral. Cada una de estas construcciones se verificó tanto mediante patrón de digestión, como mediante secuenciación Sanger.

1.9.2 Generación de la línea celular reportera K562 SEWAS84

La línea celular K562 SEWAS84 se estableció mediante la transducción de K562wt con el LV SEWAS84, y contiene en su genoma una copia del cassette SE precedido de un fragmento de 84pb del primer intrón de *WAS*, con la diana para el sistema CRISPR/Cas9 gl1 (WAS84).

El plásmido lentiviral SEWAS84 a partir del cual se generarían los LV se obtuvo incorporando el fragmento WAS84 en el plásmido lentiviral SE. Para ello, a partir de DNA genómico de K562wt, se amplificó este fragmento mediante PCR con la pareja de primers *BamHI*-WAS84 Fw y *WAS84*-*BamHI* Rv (tabla 6), que aportaron la diana *BamHI* en cada uno de sus extremos. El producto de PCR se clonó en el plásmido pCR2.1-TOPO, generando así el plásmido pCR2.1 WAS84. Posteriormente, este fragmento se clonó en el plásmido lentiviral SE, aguas arriba del promotor SFFV, tras digerir ambos plásmidos con la enzima de restricción *BamHI*.

Una vez conseguido el plásmido lentiviral SEWAS84, se produjeron LVs, y con el objetivo de obtener una línea celular que presentase una única copia del cassette SEWAS84 en su genoma, se llevó a cabo la transducción de células K562 con distintos MOIs. Transcurridas 48h desde la transducción, se analizó la expresión de eGFP en las células transducidas. Posteriormente, a partir de K562 transducidas con el MOI más bajo, en las que menos de un 10% de las células expresaban eGFP, se separó la población eGFP positiva mediante citometría de flujo, en el citómetro FACSAria III (BectonDickinson). Esta población heterogénea de células eGFP positivas se expandió, y a partir de ella se estableció una línea isogénica (K562 SEWAS84, clon1), mediante diluciones seriadas del cultivo celular en suspensión. Esta línea sería la empleada posteriormente como línea celular reportera, para analizar eficacia y seguridad de la EG utilizando el sistema CRISPR/Cas9 gl1.

1.10 GENERACIÓN DE DONADORES PARA RH

1.10.1 Donadores para el modelo celular K562 SE- δ GF-I1-P

1.10.1.1 SEED

El plásmido lentiviral SEED, que se utilizó como DNA donador en estudios de RH llevados a cabo sobre la línea celular K562 SE- δ GF-I1-P, se obtuvo introduciendo en el plásmido lentiviral SE el cassette de expresión EF1 α -dsRED, compuesto por la secuencia codificante de la proteína roja fluorescente (dsRed) bajo el promotor del factor de elongación 1 α humano (EF1 α). Este cassette se insertó aguas abajo del cassette SE, y en la misma orientación. Para ello, EF1 α -dsRED se extrajo del plásmido pGEMT- EF1 α -dsRED mediante digestión con la enzima de restricción *KpnI*, y se clonó en el sitio único *KpnI* del plásmido lentiviral SE.

1.10.1.2 SECdR, SECdR IS2, y SECdR WI1R

Los plásmidos lentivirales SECdR, SECdRIS2 y SECdR_WI1R se generaron mediante la inclusión del cassette CMV_dsRED (CdR) en los plásmidos lentivirales SE, SEIS2, y WAS_SE. El cassette CdR contiene, bajo el promotor CMV, la secuencia codificante de dsRed con la señal de poliadenilación (pA) del gen de la hormona de crecimiento bovina (bGH), y su inclusión en los plásmidos lentivirales SE, SEIS2, y WAS_SE se llevó a cabo de forma dirigida, tras su extracción con las enzimas *KpnI* y *XhoI* del plásmido lentiviral LV CdR.

De forma previa a este clonaje, mediante digestión con la enzima de restricción *EcoRI* se liberó el fragmento de 387pb del primer intrón de *WAS* del plásmido pCR2.1 WAS I₁ 387, para su inclusión en la diana *EcoRI* del plásmido lentiviral SE, precediendo al promotor SFFV, dando lugar así al plásmido WAS_SE.

1.10.2 Donador para el modelo celular K562 SEWAS84

1.10.2.1 SdR*GFP

El plásmido pUC57 SdR*GFP, usado como DNA donador para la disrupción de eGFP en el modelo celular K562 SEWAS84, está compuesto por el cassette de expresión SFFV-dsRed (SdR) seguido por la señal de pA de bGH, y por la secuencia codificante para eGFP, incluyendo diferentes mutaciones que impiden su expresión. Este plásmido fue sintetizado por GenScript.

1.10.3 Donador para el locus *WAS* endógeno

El donador para el locus *WAS* endógeno se diseñó flanqueando el cassette SE, seguido por la señal de pA de bGH, por dos brazos de homología respecto al locus *WAS*. En concreto, el brazo de homología en 5' lo componen las 760pb de *WAS* que preceden al sitio de corte del gI₉, y que abarcan parte del primer intrón, el primer exón, y parte de la región promotora. De forma similar, el brazo de homología en 3' lo componen las 760pb que suceden al sitio de corte. Aguas arriba del brazo de homología en 5' se encuentra la secuencia diana para el mismo gRNA (gI₉), en orientación invertida. Este donador, incluido en el plásmido pUC57, fue sintetizado por GenScript.

1.11 CARACTERIZACIÓN DE EFICACIA DE LOS IDLVs

1.11.1 Perfil de formas extracromosómicas

El ratio de moléculas de DNA con 2-LTR respecto al total de moléculas episomales en células transducidas con IDLV se determinó mediante qPCR, usando diferentes parejas de primers (Tabla 1) que permiten discriminar las formas con 2-LTR del total de formas episomales. Como control interno, se emplearon primers frente al locus de la

Albúmina (hAlb). Para ello, la extracción de DNA de las células transducidas se realizó transcurridas 72h desde la transducción, mediante columnas de filtrado con membrana de sílice (QIAamp DNA Mini Kit, QIAGEN). Las qPCRs se llevaron a cabo en el termociclador Stratagene MX3005P System (Agilent Technologies), mediante el kit QuantiTect SYBRGreen PCR (QIAGEN), seleccionando el siguiente programa de amplificación: 10' a 95°C, seguidos por 40 ciclos (15" a 95°C, 60" a 60°C, 60" a 72°C), finalizando con la curva de Melting. Los datos se analizaron mediante el método $2^{-\Delta \Delta Ct}$ ²¹⁴.

1.11.2 RT-qPCR para determinación de niveles de mRNA

El RNA total se obtuvo mediante el método de extracción fenólica con Trizol (Invitrogen), siguiendo las recomendaciones de la casa comercial. Esta extracción se realizó en células transducidas con IDLV, transcurridas 72h desde la transducción. La retrotranscripción de las muestras de RNA se llevó a cabo usando el kit Superscript first-strand system (Invitrogen), y una vez obtenido el cDNA, se realizó la correspondiente qPCR con el kit QuantiTect SYBRGreen PCR (QIAGEN), en el termociclador Stratagene MX3005P System (Agilent Technologies). Los primers que se usaron se muestran en la tabla 1.

1.11.3 FISH

La determinación de la localización de los genomas virales en el núcleo de células transducidas con IDLV se llevó a cabo mediante FISH, empleando una sonda Orange 8 (Alexa Fluor 555) fluorescente. Para generar la sonda SE-FISH, el DNA del plásmido vector se marcó directamente por nick translation, según las especificaciones del fabricante (Invitrogen). Las células fijadas con la solución de Carnoy se hibridaron durante la noche a 37°C con la sonda SE. Después de los lavados post-hibridación, para

la preparación de las muestras para los análisis de microscopía se usó medio de montaje SlowFade Gold Antifade Mountant con DAPI (Invitrogen). Finalmente, las imágenes fueron adquiridas en un microscopio confocal Zeiss LSM 710 (Carl Zeiss). El grado de colocación se analizó mediante el coeficiente de superposición de Manders².

1.12 CARACTERIZACIÓN DE LA EFICACIA Y ESPECIFICIDAD DE CORTE DE LAS NE

1.12.1 Cuantificación de eficiencia de corte en sitio diana

La eficiencia de corte del sistema CRISPR/Cas9 en el locus *WAS* se determinó mediante el ensayo de la T7 endonucleasa I (New England Biolabs), enzima capaz de reconocer despareamientos en el DNA, y escindir cualquiera de los tres enlaces fosfodiéster inmediatamente aguas arriba del despareamiento. A partir de DNA genómico procedente, tanto de una población mixta de células editadas, como de células no editadas como control, se amplificó mediante PCR con la pareja de primers hWAS5 Fw y hWAS5 Rv (Tabla 1) un fragmento de 2Kb del locus *WAS*, conteniendo la secuencia diana. El producto de PCR se purificó mediante columnas de filtrado con membrana de sílice para productos de PCR de más de 100pb (QIAquick PCR Purification Kit, QIAGEN), y se llevó a una concentración de 10ng/ μ l. Para la formación de heterodúplex, 180ng de producto de PCR se mezclaron con 2 μ l del buffer de la T7 endonucleasa I, en un volumen final de 20 μ l. La desnaturalización y rehibridación de la mezcla de productos de PCR para la formación de heterodúplex se llevó a cabo en el termociclador Veriti 96-Well (Applied Biosystems), mediante el siguiente programa: 95°C durante 10 min, seguidos de una rampa decreciente de temperatura desde 95°C hasta 85°C en un incremento de -2°C/s, y una segunda rampa decreciente de temperatura desde 85°C hasta 25°C, en un incremento de -0.1°C/s, finalizando a 4°C.

Tras la desnaturalización y rehibridación, los productos se digirieron con T7 endonucleasa I, durante 1h a 37°C, y se cargaron posteriormente en un gel de agarosa al 2%. La eficiencia de corte del sistema CRISPR/Cas9 se determinó mediante densitometría de bandas, aplicando la siguiente fórmula:

$$\% \text{ Corte} = 100x \left[1 - \left(1 - \frac{b + c}{a + b + c} \right)^{1/2} \right]$$

Donde **a** hace referencia a la intensidad de la banda original, no digerida, y **b + c** a la intensidad de los productos de digestión²¹⁵.

1.12.2 Análisis de efectos *off target*

La predicción de posibles sitios *off target* para nuestros sistemas CRISPR/Cas9_gRNA se llevó a cabo, en el caso del gRNA9, mediante BLAST, utilizando los parámetros predefinidos por el algoritmo de alineamiento. Para ello, teniendo en cuenta que la región del crRNA más determinante en la especificidad del sistema la configuran los 6-11 nucleótidos (nt) que preceden a la secuencia PAM³, solo se consideraron los 16 nt inmediatamente aguas arriba de la secuencia PAM, junto a esta. En el caso del gRNA1 frente al primer exón de *WAS* (gE₁), esta predicción se realizó mediante la herramienta online Cas-OFFinder, considerando toda la secuencia del crRNA, y permitiendo el despareamiento de hasta dos bases, así como la formación de pequeños bucles tanto en el DNA como en el crRNA.

Para cada uno de los gRNAs seleccionamos 6 de los potenciales sitios *off target*, y analizamos la generación de indels mediante el algoritmo ICE⁴, que permite medir la discordancia entre dos secuencias Sanger. Con el fin de secuenciar los sitios predichos

en diferentes muestras editadas, así como en muestras sin editar que sirviesen como control, se diseñaron parejas de primers para cada uno de estos sitios (Tablas 4 y 5)

Tabla 4. Potenciales sitios *off target* para el sistema CRISPR/Cas9_gE1, dirigido frente al primer exón de *WAS*. Marcados en rojo se muestran los desapareamientos con respecto a la secuencia original. En azul se muestra la correspondiente secuencia PAM. En la tabla se muestran también los primers usados para la amplificación y posterior secuenciación de cada uno de los sitios *off target*.

		Locus	Localización Cromosómica	Tipo	Primers 5' → 3'
gE ₁	crRNA	<i>WAS</i>	Xpl1.23	Gen <i>WAS</i> . Región exónica.	Fw: AAGCACTCAGCATAGGCGTGG Rv: AAGTCAAGTCAGGGGATGTT
	Secuencia diana				
E ₁ _0T1	crRNA	<i>FRG1-DT</i>	4q35.2	Gen <i>FRG1-DT</i> . Región intrónica	Fw: TTGGGGGAGAGGACAGTTT Rv: ACGTTTCCGGTTCCTTGG
	Secuencia diana				
E ₁ _0T2	crRNA	NC_000002.12	2q14.3	Región intergénica	Fw: GGCAGACACTTACTTGGCT Rv: ATGCCCTGTTCACTGCTCA
	Secuencia diana				
E ₁ _0T3	crRNA	NC_000010.11	10q26.3	Región intergénica	Fw: TGGGACAGCATTTGACCACTA Rv: GAAAAGGGAGAGTTTCCGGGT
	Secuencia diana				
E ₁ _0T4	crRNA	NC_000008.11	8q24.3	Región intergénica	Fw: CTGACCTTAATCTGTGAATGG Rv: GTGGCAGATACCACTTC
	Secuencia diana				
E ₁ _0T5	crRNA	<i>CYBBN</i>	7q33	Gen <i>CYBBN</i> . Región intrónica	Fw: GCTTCTAGGCTGGCTGGAG Rv: AIGTTGAGGCTGGGTTTTCC
	Secuencia diana				
E ₁ _0T6	crRNA	<i>CASZ1</i>	1p36.22	Gen <i>CASZ1</i> . Región intrónica	Fw: GGGTGGCTTGGGATCTATT Rv: KGAAGTACCCATGAGMAGGC
	Secuencia diana				

Tabla 5. Potenciales sitios *off target* para el sistema CRISPR/Cas9_gI9, dirigido frente al primer intrón de *WAS*. Marcados en rojo se muestran los desapareamientos con respecto a la secuencia original. En azul se muestra la correspondiente secuencia PAM. En la tabla se muestran también los primers usados para la amplificación y posterior secuenciación de cada uno de los sitios *off target*.

		Locus	Localización Cromosómica	Tipo	Primers 5' → 3'
gI_9	crRNA	GAGGCGAGGAMGGATCCAGGTC	Xp11.23	Gen <i>WAS</i> . Region intrónica.	Fw: AAGCACTCAGTCAATAGCGGTGG
	Secuencia diana	GAGGCGAGGAMGGATCCAGGTC			Rv: AAGTTGAGGTGAGGGGATTTG
I_011	crRNA	-----CAAGGACCAAGTC	3q22.3	Gen <i>ESY73</i> . Region intrónica.	Fw: AGGCGGGCTGGGAAAGCTTA
	Secuencia diana	CAGAGGATCAAGGATCAGGTCCTGG			Rv: CCCGAAGGATCAGCTCATG
I_012	crRNA	-----CAAGGACCAAGTC	5q34-q35.1	Gen <i>SLIT3</i> . Region intrónica.	Fw: CAGGAGAGTGGTTGACCAAT
	Secuencia diana	TCTCTGGCAGGACAGGTCCTGG			Rv: AAGAGTCTTACGCCAAGAGG
I_013	crRNA	-----GACCTGGTCTTGG	7q31.31	Gen <i>ANKRD7</i> . Region intrónica.	Fw: GAAACATCCAAATGAAAGG
	Secuencia diana	TGGCCGAGACCTGGTCTTGG			Rv: ATCTGTGCCAGGCATTATT
I_014	crRNA	-----GACCTGGTCTTGG	3p21.31	Gen <i>CACNA2D2</i> . Region intrónica.	Fw: CTTCTCTGAGTTGGCTTAGG
	Secuencia diana	GAGTCCAGACTCTGGTCTTGG			Rv: TATGACTCCAGGAGTACAG
I_015	crRNA	-----AAGGACCAAGTC	6p21.31	Gen <i>ABMK12</i> . Region intrónica.	Fw: GCAGATCACTTGGACCGTA
	Secuencia diana	AGCTATAGAAAGGACAGGTCCTGG			Rv: GGCCAAACCCAAATGAGCAAT
I_016	crRNA	-----GAAGGAGGCAAGTC	3p21.31	Gen <i>CTBS61D2</i> . Region intrónica.	Fw: CTCCTGGACTTACTGTCTA
	Secuencia diana	GGGAAACGAAAGGACAGGTCCTGG			Rv: TTCGGTGAATCTGACTAAGC
I_017	crRNA	-----GACCTGGTCTTGG	17q25.3	Gen <i>SEPT9</i> . Region intrónica.	Fw: GATTTGATCTGAGGAGTGA
	Secuencia diana	CTCTCCAGACTGCTCTCTTGG			Rv: CACCAGTTTCGACAGTGT

1.12.3 Tinción intracelular de WASP

El perfil de expresión de WASP en HSCs editadas se analizó mediante tinción intracelular de WASP. Para ello, se lavaron las células con FACS buffer (PBS con 2% de FBS y 2mM de EDTA) durante 5 minutos a 300g, y se fijaron con PFA al 2% en PBS, durante 20 minutos a RT. Posteriormente, las células se lavaron con PBS, y se permeabilizaron con Perm/Wash buffer (PBS con 3% de BSA y 0.2% de saponina), con una incubación de 10 minutos. Tras la permeabilización, se realizó un nuevo lavado, seguido del bloqueo de los receptores Fc, incubando las muestras con FcR blocking (Milteny), durante 15 minutos en hielo. A continuación, las muestras se incubaron durante 1h, en hielo, con el anticuerpo monoclonal anti-WASP, conseguido en conejo (abcam, EP2541Y), o con el control de isotipo (abcam, Rabbit IgG, monoclonal, EPR25A), en una dilución 1:50 en Perm/Wash buffer con un 5% de suero de cabra, para intensificar el bloqueo de los receptores Fc. Transcurrido este tiempo, se lavaron las células y se incubaron con el anticuerpo secundario APC Goat Anti-Rabbit (IgG) (abcam), durante 30 minutos en hielo. Finalmente, las muestras se lavaron una última vez, y se procedió a evaluar la expresión de WASP mediante citometría de flujo.

1.12.4 Análisis estadísticos

Todos los datos se representan como medios \pm SEM. El análisis estadístico se realizó con el software GraphPad Prism (GraphPad Software; <https://www.graphpad.com>) aplicando la prueba t no pareada de dos colas. La significación estadística se definió con un valor de $p < 0,05$. Cuando no fue posible recurrir a pruebas paramétricas, se empleó la prueba U de Mann-Whitney para la comparación de dos muestras independientes ($p < 0,05$), y el método de Dunn para comparaciones múltiples ($p < 0,05$).

Tabla 6. Primers usados en esta memoria

Primer	Secuencia	Finalidad
hWASP2 Fw	5' AGGGTTCCAATCTGATGGCG 3'	Fragmento del gen <i>WAS</i>
hWASP2 Rv	5' TTGAGAACTGGCTTGCAAGTCC 3'	
hWASP-I ₁ <i>Pfo</i> Fw	5' TCCTGGACAGGACCACGAGAAC 3'	Fragmento de 387pb del I ₁ de <i>WAS</i>
hWASP-I ₁ <i>Pfo</i> Rv	5' TCCAGGACAGCGCCAGGTACAG 3'	
<i>Bam</i> HI-WAS84 Fw	5' GGATCCATCCTCCGCTCCTCCTTCC 3'	Fragmento de 84pb del I ₁ de <i>WAS</i>
WAS84- <i>Bam</i> HI/Rv	5' GGATCCATCCTCCTGGGAAGGGTGGATT 3'	
<i>Bam</i> HI- δ GFP Fw	5' GGATCCTGAGCAAGGGCGA 3'	δ GFP
<i>Xho</i> I-GFP Rv	5' CCCTCGAGGTCGACTCTAGAGTC 3'	
qhAlb Fw	5' GCTGTCATCTTGTGGGCTGT 3'	Normalización qPCR
qhAlb Rv	5' ACTCATGGGAGCTGCTGGTTC 3'	
Δ U3 Fw	5' TGTGTGCCCGTCTGTTGTGT 3'	Genomas virales totales
PBS Rv	5' GAGTCCTGCGTCGAGAGAGC 3'	
q2LTR Fw	5' GCCTCAATAAAGCTTGCCTTG 3'	Círculos episomales con 2-LTR
q2LTR Rv	5' TGGGAGTGAATTAGCCCTTCCA 3'	
g9_Fw	5' CCAGACCTGGTCCTCCT 3'	Donadores para el locus <i>WAS</i> (Figura 45)
WAS_I_Rv	5' ATATGGAACTCTGGGGTCT 3'	
WASp_Fw	5' ATGACTACTCCTGCCACAG 3'	

RESULTADOS

1.13 OBJETIVO 1: MEJORA DE LA EFICACIA DE EDICIÓN GENÓMICA EN CÉLULAS T Y HSCs

1.13.1 Evaluación de la eficiencia y especificidad de corte en el locus *WAS* utilizando diferentes métodos de entrega del sistema CRISPR/Cas9 en células K562

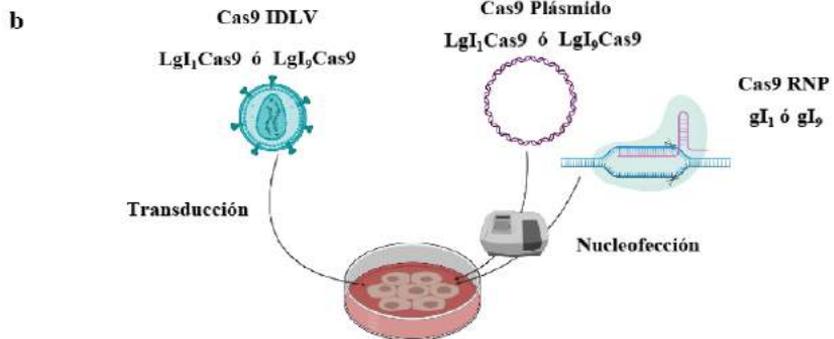
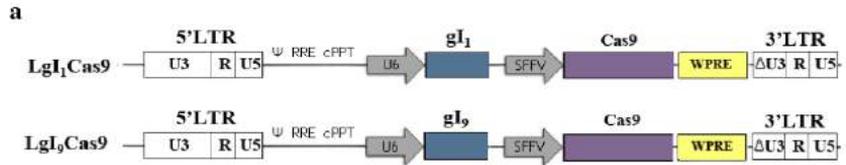
Los sistemas CRISPR/Cas9 utilizados en esta memoria, dirigidos frente al primer intrón del gen *WAS*, fueron diseñados en nuestro laboratorio, basándonos en los trabajos de Charpentier y Doudna (2012)²¹⁶, y Zhang (2013)⁸⁸. El cDNA de la nucleasa Cas9 se obtuvo mediante síntesis génica (GenScript), tras la optimización del uso de codones de la secuencia original de Cas9 de *Streptococcus pyogenes*, para favorecer una mejor expresión en células humanas. Además, se incorporó a la secuencia el epítipo HA²¹⁷, y dos señales de localización nuclear (NLSs), para favorecer la importación al interior del núcleo (coCas9). Los gRNAs frente al primer intrón de *WAS* (gl₁ y gl₉) se diseñaron buscando en dicha región la secuencia del motivo adyacente al protoespaciador (PAM) (secuencia NGG), precedida de al menos 15 nucleótidos que, en su conjunto, formaran una secuencia que no existiera en otras regiones del genoma. Para ello, las secuencias candidatas se analizaron mediante el algoritmo de alineamiento BLAST. Una vez seleccionados, se hizo un análisis de idoneidad de ambos gRNAs, utilizando el programa CRISPOR (Figura 7, c). Estos análisis predecían niveles similares de eficacia de ambos gRNAs, aunque un mayor número de posibles sitios *off-target* para el gl₉.

A continuación, se diseñaron dos cassettes de expresión, uno basado en el promotor U6 para la expresión de cada gRNA, y otro basado en el promotor SFFV, para la expresión de coCas9. Ambos cassettes, SFFV-coCas9 (SCas9) y U6- gl₁ o gl₉, fueron incluidos en un único plásmido lentiviral, obteniéndose así Lgl₁Cas9 y Lgl₉Cas9 (Figura

7, a), que permiten la entrega simultánea de los sistemas CRISPR/Cas9 específicos para el intrón 1 de *WAS* a las células diana. Esta construcción permite, además, una comparación directa de la eficacia de edición utilizando diferentes sistemas de entrega del mismo material (como plásmido o como IDLV).

A continuación, se procedió a realizar una comparativa de la eficacia de tres sistemas de entrega diferentes; dos sistemas que utilizan las construcciones Lgl₁Cas9 o Lgl₉Cas9 como base (entrega como plásmido o como IDLV), y otro que utiliza ribonucleoproteínas preformadas *in vitro* (RNP) (Figura 7, b). Como se indica en la Figura 7,b, el plásmido y la ribonucleoproteína se transfieren por nucleofección, dado que es el método más eficaz de transferencia, tanto en nuestras células diana finales (HSCs) como en nuestras células modelo (K562, línea celular de linaje hematopoyético idónea para modelar WAS). El otro sistema utiliza vectores lentivirales deficientes para integración (IDLVs), con la finalidad de utilizar las propiedades de estos vectores para conseguir niveles óptimos y transitorios de expresión de Cas9 y gRNA en las células diana, en función del MOI utilizado y del diseño del vector. En la Figura 7, d se muestra un ejemplo representativo de las eficacias de corte obtenidas en K562 mediante los diferentes sistemas, utilizando los dos gRNAs seleccionados.

Con todos los sistemas se obtuvieron eficiencias similares, aunque el sistema de nucleofección parece ser el más robusto, sobre todo cuando se utilizan las últimas actualizaciones en cuanto a protocolo de nucleofección y concentraciones de RNP. Además, se pudo observar una elevada variabilidad de la eficiencia de ambos gRNAs en función del sistema de transferencia. Así, el gl₁ era el más eficaz como RNP en K562, mientras que el gl₉ ofrecía mejores resultados como plásmido.



c

gRNA	Secuencia	Predicción de especificidad		Pred Eficacia	Cambios en el marco de lectura		Predicción sitios <i>off target</i>	
		MIT Score	CFD score	Reglas de Doench	Debido a deleciones	Deido a Indels	Totales	CFD score
1	CGTCATAATC CACCCITCCC	75	89	56	73	77	205	0 con score > 1.0
9	CAGGCAGGAA GGACCAGGTC	48	72	52	57	73	682	11 con score > 1.0

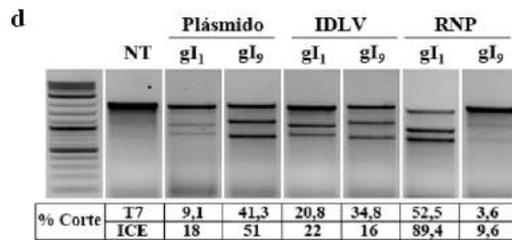


Figura 7. Eficiencia de Sistemas CRISPR/Cas9 específicos para *WAS* en K562, suministrados mediante distintos métodos de entrega. a) Esquema de los plásmidos lentivirales Lgl₁Cas9 y Lgl₉Cas9, expresando coCas9 optimizada bajo el promotor SFFV (SCas9), y el correspondiente gRNA bajo el promotor U6. b) Esquema de los procedimientos empleados para vehicular los sistemas CRISPR/Cas9 específicos para *WAS* al interior de K562. c) Análisis de idoneidad de los diferentes gRNAs dirigidos frente al primer intrón de *WAS* (gl₁ y gl₉), realizado *in silico*, con el programa CRISPOR (<http://crispor.tefor.net/>). La tabla muestra la secuencia de cada uno de los gRNAs, así como una predicción de su especificidad, eficiencia, y potenciales sitios *off target*, en base a distintos algoritmos (ver material y métodos). d) Ejemplo representativo de la eficiencia de corte de los sistemas CRISPR/Cas9 dirigidos frente al primer intrón del locus *WAS*, determinada mediante el ensayo de la T7, y mediante el algoritmo ICE (ver material y métodos).

Se ha descrito previamente que la concentración de Cas9 y/o gRNA pueden afectar la especificidad de corte, incrementando el número y frecuencia de cortes fuera del sitio diana (sitios *off-target*)²¹⁸⁻²²⁰. Partiendo de la hipótesis de que los tres sistemas de transferencia van a ofrecer diferentes niveles de expresión de Cas9, se procedió a determinar si la generación de indels en sitios *off target* podría seguir un patrón distinto en función del método escogido para hacer llegar el sistema CRISPR/Cas9 a las células diana. Para ello, se seleccionaron 7 posibles sitios *off target* (I₁_OT1 - I₁_OT7) (°Tabla 5) para el sistema CRISPR/Cas9 gl₉ (que ofrece un mayor número de sitios *off-target* posibles, en base al análisis de idoneidad con CRISPOR), y se evaluó corte en cada uno ellos al utilizar los tres sistemas de entrega (Figura 8).

Con la finalidad de poder comparar los diferentes sistemas de entrega, se utilizaron muestras en las que se había obtenido una eficiencia similar de corte en la diana (entre 8 - 20% de corte) para comparar los niveles de indels en los sitios I₁_OT1 - I₁_OT7. Estos sitios se secuenciaron mediante secuenciación Sanger, y a partir de estas secuencias, se determinó el porcentaje de indels en cada caso mediante el algoritmo ICE. El sitio

I₁_OT6 se excluyó de los análisis, debido a que esta cuantificación no pudo llevarse a cabo por la mala calidad de las secuencias obtenidas.

Considerando el conjunto de los sitios *off target* analizados, no se observaron diferencias significativas en el porcentaje global de indels en sitios *off target* en función del método empleado para la entrega del sistema CRISPR/Cas9 g₉ (Figura 8). No obstante, para poder concluir que no existe una variación en los efectos *off target* globales del sistema CRISPR/Cas9 g₉ en función del método de entrega, sería necesario realizar un análisis no sesgado.

Por otro lado, la consideración de cada uno de los sitios *off target* de forma individual (Figura 8, b) si sugiere un patrón de generación de indels diferente en función del método de entrega. Por ejemplo, Lg₉Cas9 IDLV no generó indels en I₁_OT1, en contraposición a Lg₉Cas9 plásmido y RNP g₉, pero fue el único método de entrega con el que se detectaron indels en I₁_OT2.

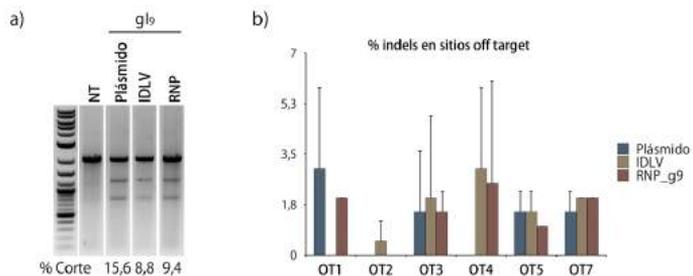


Figura 8. Sistemas CRISPR/Cas9 específicos para *WAS*, suministrados a K562 mediante plásmido, IDLV o RNP, consiguen editar *WAS* de forma eficiente, presentando efectos *off target* similares para eficiencias de corte en la diana equivalentes. a) Experimento representativo del análisis de la eficiencia de corte del sistema CRISPR/Cas9 gI_9 suministrado a K562 mediante plásmido, IDLV o RNP. El % de corte en cada caso se estimó mediante el ensayo de la T7. b) Porcentaje de indels generados en potenciales sitios *off target* por el gI_9 ($I_1_OT1 - I_1_OT7$) analizados con el algoritmo ICE. Los datos representan la media \pm el error estándar de al menos 2 experimentos independientes.

1.13.2 Evaluación de la eficiencia y especificidad de corte en el locus *WAS* utilizando diferentes métodos de entrega del sistema CRISPR/Cas9 en células T y HSCs

Diferentes tipos celulares tienen resistencias diferentes a ser modificadas genéticamente. Efectivamente, no todo lo que observamos en las células K562 va a ser trasladable a células HSCs y/o células T. Nuestro objetivo inmediato fue, por tanto, estudiar la eficacia y especificidad de los sistemas desarrollados en nuestras células diana.

A pesar de que en K562 obtuvimos eficiencias de corte similares con los tres métodos de entrega, estos resultados no se repitieron en células T y HSCs, donde RNP fue el único método que proporcionó eficiencias de corte consistentes.

En el caso de células T, se emplearon los tres métodos de entrega para evaluar la eficacia de corte, tanto del sistema CRISPR/Cas9 gI_1 , como del sistema CRISPR/Cas9 gI_9 . El algoritmo ICE nos permitió determinar la eficiencia, y comprobar que la entrega de ambos sistemas en forma de plásmido, o de IDLV, no generaba indels en células T (datos no mostrados), mientras que al entregarlos como RNP se obtenían eficiencias

elevadas (superior al 50% en el caso de RNP gl_1 , y próximas al 90% en el caso de RNP gl_9) (Figura 9). También pudimos observar diferencias en el patrón de indels generados por uno y otro gRNA en el locus *WAS*. En el caso de RNP gl_9 (Figura 9, gráfica derecha), el indel más frecuente consistió en la pérdida de un único nucleótido. Este tipo de indels se ha descrito como uno de los generados con mayor frecuencia por sistemas CRISPR/Cas9¹⁰⁹. En el caso de RNP gl_1 (Figura 9, izquierda), la pérdida de un único nucleótido fue, también, el tipo de indel mayoritario. Sin embargo, este gRNA generó, adicionalmente, otro tipo de indels con frecuencias relativamente elevadas (por ejemplo, la pérdida de 5 nucleótidos, en este caso, con una frecuencia superior al 10%).

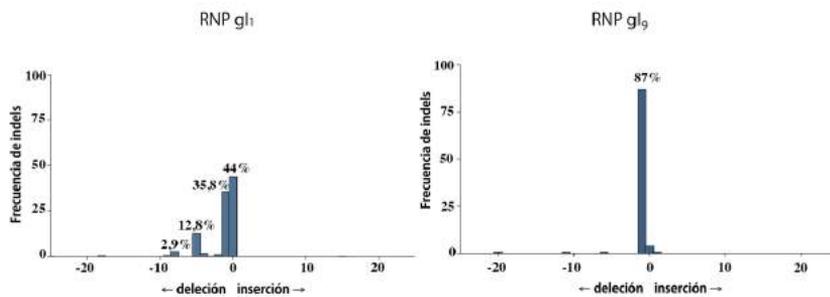


Figura 9. Eficiencia en células T de Sistemas CRISPR/Cas9 específicos para *WAS*, suministrados mediante RNP. Gráfica mostrando el perfil de indels generados en células T nucleofectadas con RNP gl_1 (izquierda) ó RNP gl_9 (derecha), determinado mediante el algoritmo ICE. La coordenada 0 representa el sitio de corte, valores negativos representan deleciones de distinta longitud, y valores positivos, inserciones.

En HSCs seleccionamos únicamente el sistema CRISPR/Cas9 gl_9 , para analizar su eficiencia de corte en el primer intrón de *WAS* al ser vehiculizado hasta el interior de

estas células con los diferentes sistemas. En este caso, la eficiencia de corte se determinó tanto a partir del ensayo de la T7, como mediante el algoritmo ICE (Figura 10, a). Tal y como sucedió en células T, RNP fue el único método de entrega con el que se obtuvo corte en el locus *WAS* en HSCs. En este sentido, pudimos apreciar que la eficiencia de corte reflejada mediante T7 era muy inferior a la que nos indicaba el algoritmo ICE. Esto podría deberse a que, tal y como sucedía en células T, el tipo de indel mayoritario generado por RNP gl₉ consistió en la pérdida de un único nucleótido (Figura 10, b). El hecho de que una misma mutación fuese tan frecuente iría en detrimento de su detección mediante el ensayo de la T7. Por otro lado, se ha descrito que la T7 no siempre resuelve eficientemente desapareamientos de una única base. Ambos factores podrían estar influyendo, por tanto, en que mediante la T7 se estuviesen subestimando las eficiencias de corte de nuestros sistemas de EG, representando el algoritmo ICE un sistema más robusto.

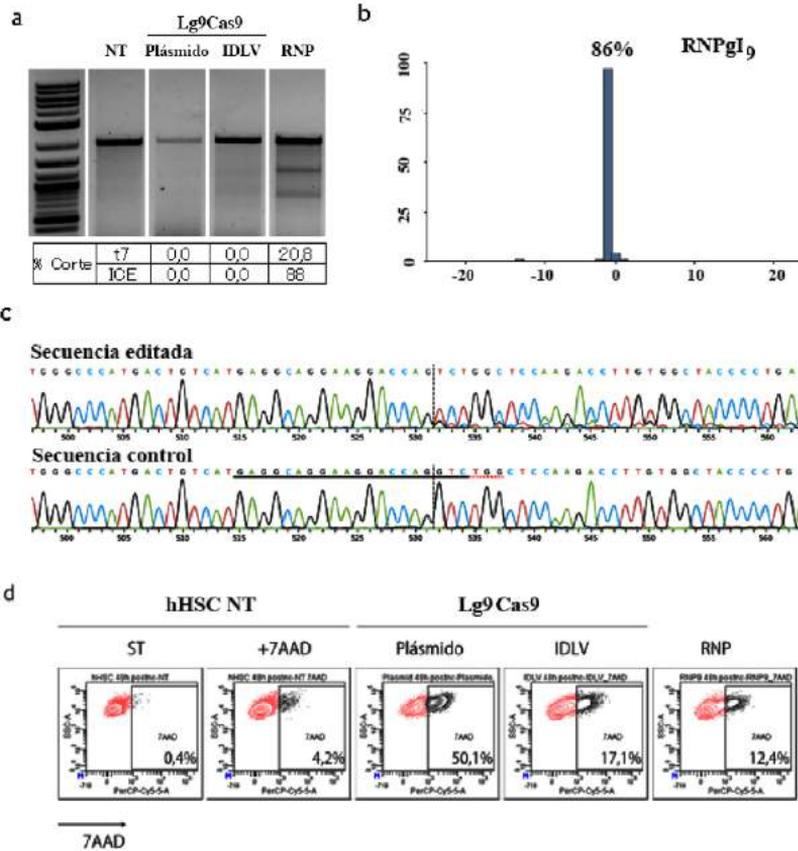


Figura 10. Eficiencia en HSCs de Sistemas CRISPR/Cas9 específicos para *WAS*, suministrados mediante distintos métodos de entrega. a) Análisis de la eficiencia de corte del sistema CRISPR/Cas9 gl₉ suministrado a HSCs en d6 post-aislamiento, mediante plásmido (nucleofección), IDLV (transducción), y RNP (nucleofección). El % de corte en cada caso se estimó mediante el ensayo de la T7, así como mediante el algoritmo ICE. b) Gráfica mostrando el perfil de indels generados en HSCs nucleofectadas con RNP, determinado mediante el algoritmo ICE. La coordenada 0 representa el sitio de corte, valores negativos representan deleciones de distinta longitud, y valores positivos, inserciones. c) Cromatograma procedente de población de HSCs editadas (superior) enfrenteado frente a cromatograma a partir de población original de HSCs, no editada (inferior). Subrayada,

aparece la secuencia diana, y delimitado por una línea transversal aparece el punto de corte. Aguas abajo del punto de corte se observa un solapamiento de lecturas en la muestra editada, que no aparece en la población original. **d)** Análisis de la viabilidad celular en función del método de entrega. Este análisis se llevó a cabo mediante citometría de flujo, marcando las células con 7AAD, 48h tras la entrega del sistema CRISPR/Cas9.

En HSCs se evaluó, además, cómo afectaba a la viabilidad celular la entrega del sistema CRISPR/Cas9 mediante cada uno de los métodos de entrega. Para ello, 48h tras la entrega del sistema CRISPR/Cas9 *gl*₉, las células se marcaron con 7AAD, y se midió viabilidad celular mediante citometría de flujo. Pudo observarse que el sistema de entrega que afectaba en mayor medida a la viabilidad celular era la entrega en forma de plásmido (Figura 10, d). La entrega del sistema CRISPR/Cas9 *gl*₉ en forma de RNP fue, además de la única que permitió editar el locus *WAS* en HSCs, la que afectó en menor medida la viabilidad celular de HSCs. Por tanto, en relación con los datos observados, RNP constituye, hasta el momento, el método de entrega más robusto para la edición, tanto de células T, como de HSCs.

1.13.3 Desarrollo de nuevos sistemas de entrega más eficaces basados en IDLVs.

A pesar de los desalentadores datos de la eficacia de los IDLVs en células T y HSCs mostrados anteriormente, seguíamos teniendo la hipótesis de que incrementando los niveles de expresión y/o título de los IDLVs podríamos lograr los niveles necesarios de Cas9 y gRNA para editar de forma eficaz y segura estas células. Uno de los argumentos a favor de continuar con esta idea era que podríamos conseguir llegar a los niveles justos (threshold) de Cas9 y/o gRNA necesarios para lograr el corte. Este método de

entrega proporcionando la cantidad necesaria y suficiente del sistema CRISPR/Cas9 podría tener una repercusión en la especificidad y toxicidad del sistema, reduciendo potenciales sitios *off-target* y mejorando la viabilidad de las células diana. Para ello, necesitamos mejorar los IDLVs para conseguir la expresión de Cas9 y gRNA a niveles que se encuentren alrededor del umbral requerido para que Cas9 corte. Por otro lado, el desarrollo de estos IDLVs mejorados, no solo sería de aplicación a la EG, sino a otros casos en que se requiera expresión transitoria y/o no integrativa. Por lo tanto, nuestro siguiente objetivo fue mejorar los IDLVs como herramienta de entrega.

Persiguiendo este objetivo, nos propusimos incluir en el esqueleto de IDLV el elemento aislador IS2, diseñado previamente en nuestro laboratorio¹⁷², y compuesto por un elemento SAR sintético (SAR2), y por el aislador HS4-650, derivado del aislador cHS4 del locus de la globina de pollo (Figura 11). Este elemento combina la capacidad de mejorar los niveles de expresión, por la disposición de las secuencias SAR a establecer interacciones con la matriz nuclear, con la función aisladora del aislador HS4, y su inclusión en vectores lentivirales integrativos (LVs) demostró prevenir el silenciamiento, reducir la variabilidad de la expresión y mejorar los niveles de expresión en células especialmente relevantes, como son células madre embrionarias humanas (hESC) indiferenciadas y células hematopoyéticas derivadas de hESC. Sobre la base de estos datos previos, consideramos que la inclusión de IS2 en IDLV podría tener efectos similares sobre sus niveles de expresión, lo que podría facilitar alcanzar el *threshold* requerido para lograr corte en las células diana, al utilizar IDLV como método de entrega del sistema CRISPR/Cas9.

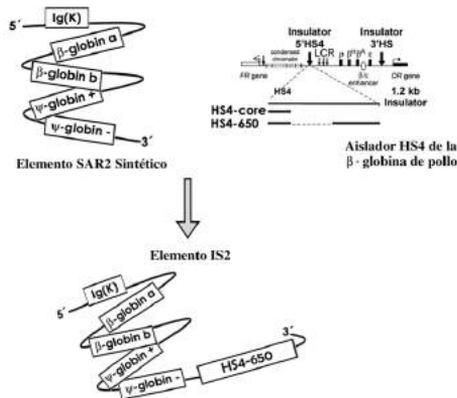


Figura 11. Esquema representativo del elemento aislador IS2 (abajo), compuesto por el elemento SAR2 sintético (izquierda), conteniendo 5 regiones MAR/SARs de anclaje a la matriz nuclear, y por el fragmento aislador HS4650, derivado del aislador cHS4 del locus de la β -globina de pollo.

Antes de estudiar el posible efecto de la inclusión de IS2 en IDLVs codificando el sistema CRISPR/Cas9, decidimos comenzar estudiando su efecto en IDLVs codificando el gen reportero eGFP, que nos proporcionasen una lectura sencilla de las mejoras introducidas en los IDLVs. Para ello, incorporamos el elemento IS2 en el 3'LTR del plásmido lentiviral SE, que contiene el cassette de expresión SFFV-eGFP (SE), y en el plásmido SEWP, que adicionalmente incluye el regulador postranscripcional WPRE (WP)²²¹ (Figura 12).

La eficacia de IDLVs generados a partir de estos nuevos esqueletos lentivirales (IS2-IDLVs) se evaluó en las líneas celulares 293T y K562, en nuestras células diana, así como en otras células de interés, como pueden ser células progenitoras neuronales humanas

(NPCs) y neuronas diferenciadas (NCs), células troncales pluripotentes inducidas (iPSCs), y fibroblastos humanos.

1.13.3.1 La inclusión del elemento IS2 en el LTR de IDLVs mejora sus niveles de expresión en las líneas celulares K562 y 293T

Para potenciar el posible efecto de la inclusión de IS2 en IDLV, este elemento se incorporó en la región $\Delta U3$ del 3'LTR de los plásmidos lentivirales SE y SEWP (Figura 12), dado que una copia de la región U3 del 3'LTR pasa a formar parte del 5'LTR tras la retrotranscripción. De forma que aquellos episomas que en el núcleo de células transducidas adoptasen una configuración como círculo con 2-LTR, contendrían dos copias de IS2, lo que podría suponer un mayor efecto.

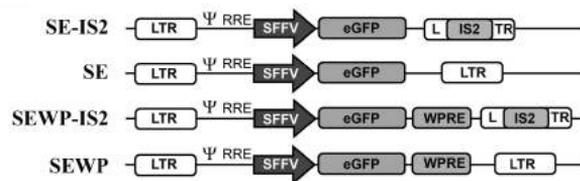


Figura 12. Representación esquemática de los vectores SE-IS2, SE, SEWP-IS2 y SEWP. eGFP, proteína de fluorescencia verde mejorada (del inglés, enhanced Green Fluorescence Protein); IS2, aislador quimérico HS4650-SAR2; LTR, repeticiones repetidas largas de HIV-1 (del inglés, Long Terminal Repeats); RRE, elemento de respuesta a Rev (del inglés Rev Responsive Element); Ψ, Señal de empaquetamiento de HIV-1; SFFV, promotor del virus formador de foco de bazo (del inglés, Spleen Focus Forming Virus); WPRE, elemento regulador postranscripcional del virus de la hepatitis Woodchuck (del inglés, Woodchuck Postranscriptional Regulatory Element).

El efecto de la inclusión de IS2 en los niveles de expresión de IDLVs se evaluó en las líneas celulares K562 y 293T, idóneas para este tipo de estudios por su alta permisividad y por su alta capacidad para expresar proteínas recombinantes²²².

Para llevar a cabo este análisis, ambos tipos celulares se transdujeron con los diferentes IDLVs, con un mismo MOI en todos los casos, estimado mediante RT-qPCR (ver material y métodos). Con el objetivo de mantener el porcentaje de células eGFP⁺ por debajo del 50%, se fijó un MOI de 0.3 partículas virales por célula. El porcentaje de células eGFP positivas y los niveles de expresión, medidos como la intensidad media de fluorescencia (MFI) de la población eGFP positiva, se determinaron mediante citometría de flujo, 3 días después de la transducción.

En K562, la inclusión del elemento IS2 en IDLVs causó un aumento en los niveles de expresión de eGFP (Figura 13, paneles inferiores), así como en el porcentaje de células que expresaban eGFP (Figura 13, paneles superiores). Este aumento, no obstante, solo fue significativo en ausencia del regulador postranscripcional WPRE.

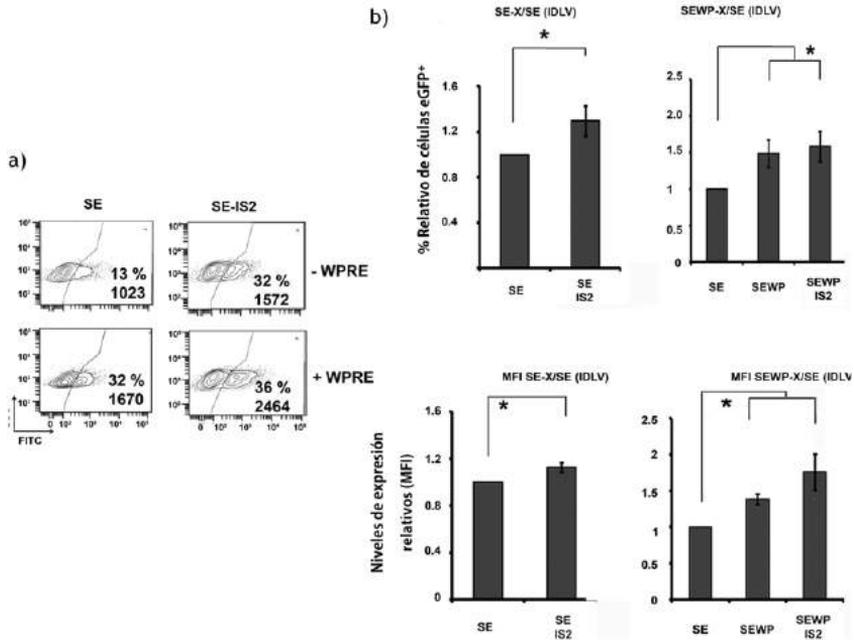


Figura 13. La inclusión del elemento IS2 en IDLVs mejora los niveles de expresión de eGFP en células K562. **a)** Plots representativos mostrando los perfiles de expresión de eGFP en K562 transducidas con los diferentes IDLVs. Los porcentajes (%) y los niveles de expresión (MFI) de cada población eGFP+ se muestran en los plots correspondientes. **b)** Gráficas mostrando los % relativos de células eGFP+ (gráficas superiores) y los niveles relativos de expresión de eGFP en K562 transducidas con SE-IS2 y SE-IDLVs, en ausencia (izquierda) o en presencia (derecha) del elemento WPRE. Para mantener el porcentaje de células eGFP+ por debajo del 50%, se usó un MOI de 0.3. La región correspondiente a la población eGFP+ se delimitó en base a K562 no transducidas, permitiendo entre un 0.2% y un 0.7% de células no transducidas dentro de esta región. Estos porcentajes residuales fueron sustraídos de los % obtenidos para los diferentes vectores y condiciones del análisis. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

En 293, la incorporación del elemento IS2 en IDLVs provocó un aumento significativo en los niveles de expresión de eGFP, tanto en ausencia, como en presencia del elemento WPRE (Figura 14, paneles inferiores). También se observó un aumento en el porcentaje de células que expresaban eGFP, aunque significativo únicamente en ausencia de WPRE (Figura 14, paneles superiores).

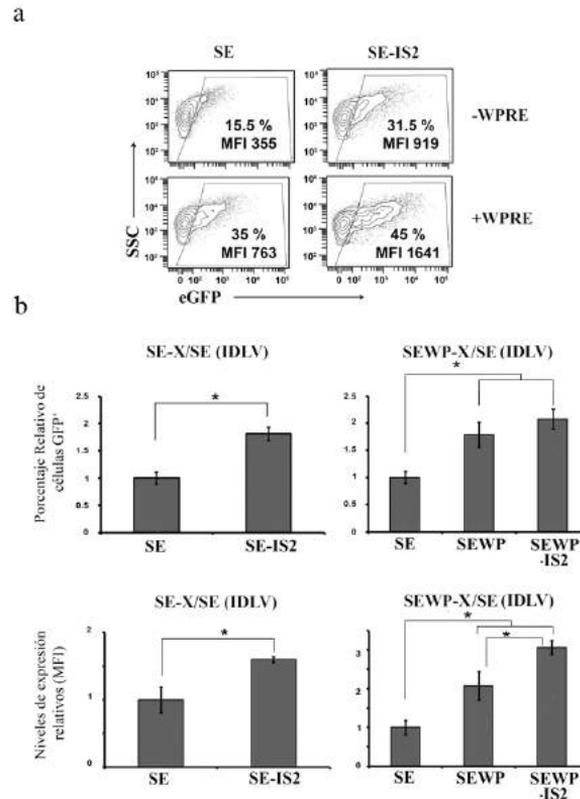


Figura 14. La inclusión del elemento IS2 en IDLVs mejora los niveles de expresión de eGFP en células 293T. a) Plots representativos mostrando los perfiles de expresión de eGFP en 293T transducidas con los diferentes IDLVs. Los porcentajes (%) y los niveles de expresión (MFI) de cada población eGFP+ se muestran en los plots correspondientes. b) Gráficos mostrando los % relativos de células eGFP+ (gráficos superiores) y los niveles relativos de expresión (gráficos inferiores) en células 293T transducidas con SE-IS2 y SE-IDLVs, en ausencia (izquierda) o en presencia (derecha) del elemento WPRE. Para mantener el porcentaje de células eGFP+ por debajo del 50%, se usó un MOI de 0.3. La región correspondiente a la población eGFP+ se delimitó en base a 293T no transducidas, permitiendo entre un 0.2% y un 0.7% de células no transducidas dentro de esta región. Estos porcentajes residuales fueron sustraídos de los % obtenidos para los diferentes vectores y condiciones del análisis. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

Para comprobar si este efecto se mantenía al incrementar el MOI, se transdujeron células 293T con MOIs crecientes de IDLVs (desde 0.1 hasta 10 partículas por célula). Mediante citometría de flujo se midieron porcentaje de células eGFP⁺ y MFI, transcurridos 3 días desde la transducción. El efecto de la inclusión del elemento IS2 en IDLVs observado al transducir 293T con un bajo MOI se mantuvo con MOIs superiores (Figura 15).

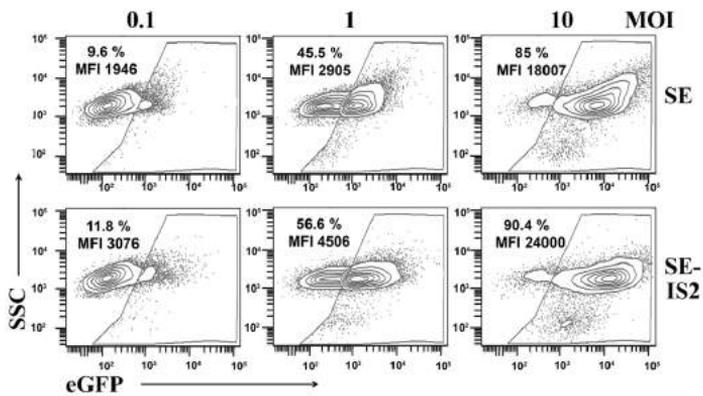


Figura 15. Efectos del elemento IS2 en IDLVs, a diferentes MOIs. Plots representativos mostrando los perfiles de expresión de eGFP en células 293T transducidas con MOIs crecientes de IDLVs SE (plots superiores) y SE-IS2 (plots inferiores). En cada plot se indica el porcentaje (%) de células eGFP⁺ y los niveles de expresión (MFI).

1.13.3.2 El elemento IS2 no previene el silenciamiento de IDLVs debido a HDAC

Los IDLV tienden a posicionarse, en el núcleo de células transducidas, en dominios represivos de la transcripción, enriquecidos en histona desacetilasas (HDAC), y como consecuencia, a sufrir silenciamiento epigenético que involucra la desacetilación de histonas, lo que lastra su potencial transcripcional^{223,224}. La prevención de la desacetilación de histonas (HDAC), el mayor mecanismo subyacente a la débil actividad transcripcional de IDLVs, podría explicar los mayores niveles de expresión alcanzados con SE-IS2-IDLVs en 293T. Con el fin de estudiar esta posibilidad, se analizaron niveles de expresión de eGFP en 293T transducidas con SE- y SE-IS2-IDLVs, en presencia y ausencia de apicidin 0.4µM, un inhibidor de desacetilación de histonas (HDACi). La adición de apicidin mejoraba los niveles de expresión en la misma medida, tanto en células transducidas con SE-IDLVs, como en las células transducidas con SE-IS2-IDLVs (del orden de 2.9 veces y 2.35 veces, respectivamente) (Figura 16), lo que sugiere que el mecanismo por el que la inclusión de IS2 en IDLVs mejora los niveles de expresión es independiente de HDAC.

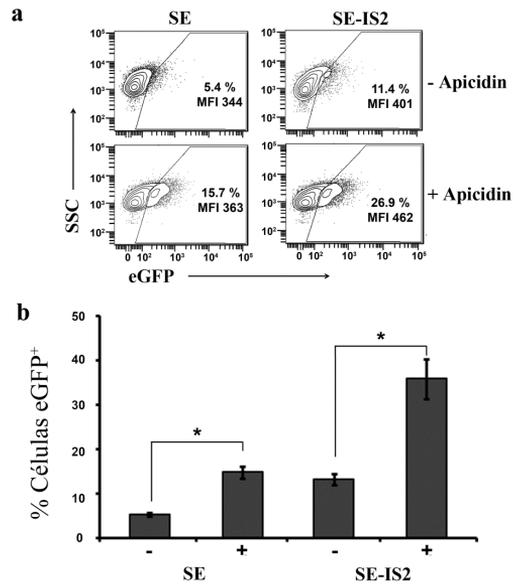


Figura 16. La expresión de eGFP en células 293T transducidas con IDLVs mejora en presencia de Apicidin, independientemente de la presencia o no del elemento IS2. a) Plots representativos mostrando los niveles de expresión de eGFP en 293T transducidas con SE o SE-IS2 IDLVs, con un MOI de 0.2, en ausencia o presencia de Apicidin 0.4 μ M. b) Gráficas mostrando el % de células eGFP⁺ en 293T transducidas con SE o SE-IS2-IDLVs, con un MOI de 0.2, en ausencia (-) o presencia (+) de Apicidin 0.4 μ M. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (*p < 0.05).

1.13.3.3 La inserción del elemento IS2 en IDLVs no afecta el empaquetamiento de RNA en las partículas virales, pero reduce la cantidad de episomas en las células diana.

La inserción de fragmentos largos en el 3'LTR de LV se ha relacionado previamente con una reducción en su eficacia. Estas inserciones no afectan a la producción de partículas virales, pero reducen la eficiencia de retrotranscripción (RT) en células diana²²⁵⁻²²⁷. Por ello, se analizó el efecto de IS2 (de 1.2Kb de longitud) en la producción de vectores, en los niveles de retrotranscripción, y en los niveles de expresión, en células 293T en este caso, por su idoneidad para este tipo de estudios, como se ha mencionado anteriormente²²².

La eficacia en la producción de vectores, medida como las UT/ml, se estimó mediante RT-qPCR (ver material y métodos). Como se esperaba, no se observó ningún efecto del elemento IS2 en la producción de LVs, ni de IDLVs.

De acuerdo con observaciones previas^{172,225,226,228}, la inserción del elemento IS2 en el 3'LTR del esqueleto lentiviral mostró un efecto negativo en los niveles de expresión en células 293T transducidas con SE-IS2-LVs (vectores integrativos) (Figura 17, a), lo que se reflejó también en una reducción en los productos de RT (Figura 17, b). Aunque la inserción de IS2 provocó una reducción similar en los productos de RT en 293T transducidas con SE-IS2-IDLVs (no integrativos) (Figura 17, b), se seguía observando una mejora significativa tanto en el porcentaje de células eGFP⁺, como en los niveles de expresión (Figura 17, a).

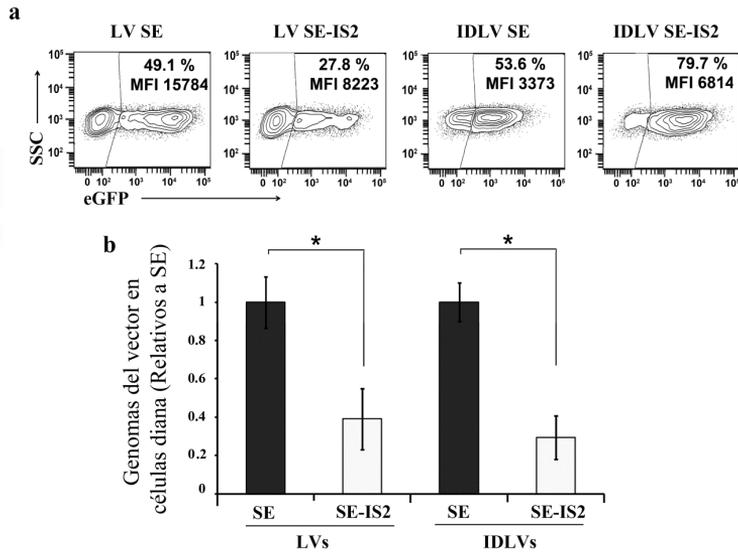


Figura 17. La inserción del elemento IS2 en el esqueleto lentiviral reduce la eficiencia de retrotranscripción en las células diana. **a)** Plots representativos mostrando los perfiles de expresión de 293T transducidas con LVs (izquierda), e IDLVs (derecha), con (SE-IS2) o sin (SE) el elemento IS2. Todos los experimentos se llevaron a cabo usando 0.7 partículas virales/célula. **b)** Gráficas mostrando la abundancia de productos de retrotranscripción (genomas del vector) en 293T transducidas con LVs(izquierda) e IDLVs(derecha), con (SE-IS2) o sin (SE) el elemento IS2, 72h después de la transducción. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

1.13.3.4 Episomas de IDLVs que contienen el elemento IS2 producen mayores niveles de RNAm, y muestran distinta localización en el núcleo

El elemento IS2 está compuesto por una secuencia SAR sintética (SAR2) y por 650pb del aislador del locus de la β -globina de pollo²²⁹. Estos elementos pueden mejorar la estabilidad episomal, lo que también puede conducir a un incremento de la expresión. Con el objetivo de esclarecer el mecanismo implicado en el efecto de la inclusión de IS2 en IDLVs, se analizó la eficacia de transcripción de los episomas SE-IS2 con respecto a los episomas SE. Para ello, se transdujeron células 293T con MOIs equivalentes de SE-IDLVs y SE-IS2-IDLVs, y se midieron los niveles de RNAm de eGFP 72h post-transducción, normalizando estos niveles con respecto a la cantidad de genomas del vector en cada caso (Figura 18, a). Este análisis mostró que episomas SE-IS2 producían 6-7 veces más RNAm que episomas SE.

Una vez se comprobó que el elemento IS2 mejoraba la transcripción de los episomas SE-IS2, se analizó si esto era debido a un aumento en la estabilidad de los mismos. Para ello, se analizaron a diferentes tiempos post-transducción (desde 24h hasta 7 días) los niveles de expresión (Figura 18, b) y la abundancia relativa de episomas SE-IS2 con respecto a episomas SE (Figura 18, c). Los datos obtenidos indican que el elemento IS2 no influye en la estabilidad de la expresión de IDLVs (Figura 18, b), ni en la estabilidad de los episomas (Figura 18, c).

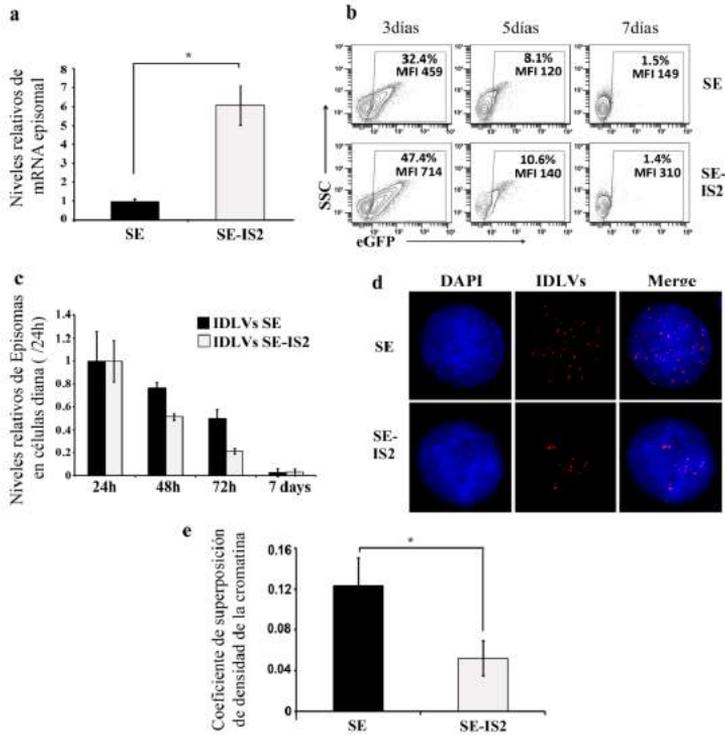


Figura 18. El elemento IS2 no afecta a la estabilidad de los episomas de IDLVs pero mejora la eficacia de transcripción, reposicionándolos en dominios del núcleo con baja densidad de cromatina. **a)** Gráfica mostrando la cantidad relativa de RNAm de eGFP en células 293T transducidas con SE-IDLVs y SE-IS2-IDLVs, normalizada respecto a la cantidad de genomas del vector en las células diana, 72h después de la transducción. **b)** Plots representativos mostrando los perfiles de expresión de eGFP en 293T transducidas con IDLVs SE (plots superiores) y SE-IS2 (plots inferiores), con un MOI de 0.5, analizados a días 3, 5 y 7 post-transducción. En cada plot se muestran los porcentajes (%) y los niveles de expresión (MFI) de la población eGFP⁺. **c)** Cantidad relativa de episomas SE-IS2-IDLVs relativos a SE-IDLVs, a diferentes tiempos post-transducción. **d)** Imágenes de confocal representativas, mostrando la distribución nuclear de los episodios correspondientes a SE y SE-IS2IDLVs. Las células fueron transducidas con un MOI de 10, con el objetivo de conseguir un número alto de episomas dentro de la célula. Las células transducidas fueron fijadas y permeabilizadas, e incubadas con una sonda frente a eGFP, marcada con Alexa Fluor 555 (IDLVs marcados en rojo). **e)** Gráficas mostrando

la colocalización entre la cromatina (tinción con DAPI) y los IDLVs, en células 293T transducidas con SE- y SE-IS2-IDLV. El grado de colocalización se determinó mediante el coeficiente de superposición de Manders. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

Tomando en consideración que los elementos SAR interaccionan en la matriz nuclear con factores de transcripción, y complejos remodeladores de la cromatina¹⁶⁵, así como la capacidad de los elementos aisladores basados en CHS4 de aumentar la eficiencia episomal, en parte, debido a la interacción entre el elemento CHS4 con la matriz nuclear a través de CTCF¹⁶⁷, decidimos comprobar si la presencia de IS2 en IDLVs podría estar favoreciendo la localización nuclear de los episomas en sitios transcripcionalmente activos (los cuales, generalmente presentan baja densidad para tinción con DAPI).

Para ello, sobre células 293T transducidas con SE-IDLVs y SE-IS2-IDLVs con un MOI de 10, se determinó la localización de los episomas en el núcleo mediante hibridación *in situ* fluorescente (FISH), empleando una sonda a partir del plásmido SE. Como se aprecia en la Figura 18, d, los episomas correspondientes a SE-IDLV se distribuyen uniformemente en el núcleo, mientras que los episomas SE-IS2-IDLV tienden a agregarse, siguiendo un patrón distinto. En base al análisis de colocalización DAPI/sonda, los episomas que contienen IS2 se localizan preferentemente en regiones con menor señal de DAPI, en comparación con episomas SE (Figura 18, d, e; Figura 19). Estos datos indican que el efecto de IS2 en episomas SE-IS2 es probablemente debido, al menos en parte, a un reposicionamiento de los episomas en regiones transcripcionalmente activas del núcleo.

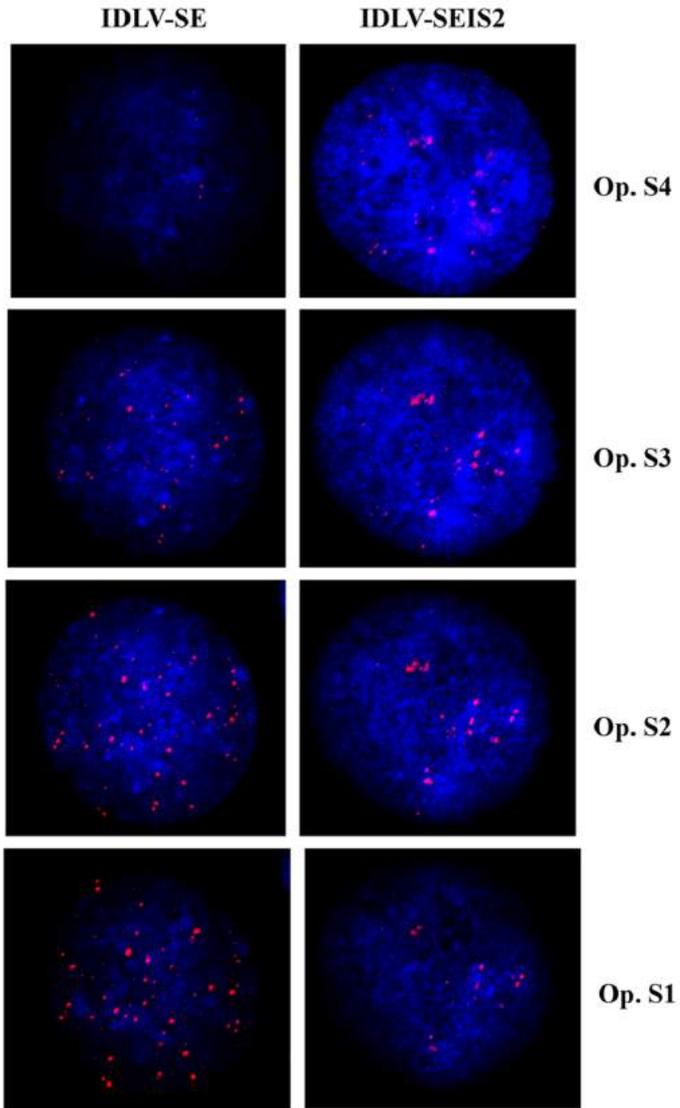


Figura 19. Imágenes de confocal mostrando la localización nuclear de los episomas del vector. Células 293T fueron transducidas con SE (izquierda) y SE-IS2 (derecha) IDLVs con un MOI de 10. Transcurridas 72h desde la transducción, las células fueron fijadas y permeabilizadas, e incubadas con una sonda frente a eGFP, marcada con Alexa Fluor 555. Las imágenes muestran la distribución de los episomas del vector en el núcleo de células transducidas, en 4 secciones ópticas (Op. S1-4).

1.13.3.5 La inserción del elemento IS2 en IDLVs reduce la formación de círculos 2-LTR, de baja expresión

La integrasa, cuyo dominio catalítico se encuentra mutado en IDLVs, no puede catalizar la integración del genoma de IDLVs en el genoma de la célula hospedadora, por lo que estas moléculas son procesadas por las vías de reparación del DNA de la célula, originando episomas circulares con 2-LTR (círculos 2-LTR), si el procesamiento se ha producido por NHEJ, o 1-LTR, si ha tenido lugar mediante RH (Figura 19).

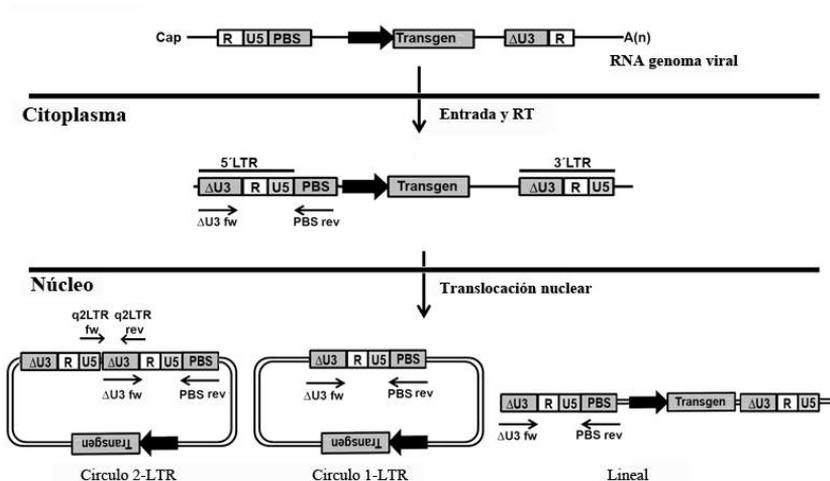


Figura 20. Esquema que muestra las diferentes formas del genoma de IDLV durante la transducción de las células diana. El RNA que constituye el genoma del vector (arriba) entra al citoplasma de la célula diana (centro), donde se produce la RT, y se generan moléculas de DNA con el 5' Δ U3RU5 LTR completo. Una vez que el complejo de preintegración entra en el núcleo, los IDLVs generan principalmente episomas circulares con 1-LTR y 2-LTR, aunque también se pueden encontrar algunas formas de DNA lineales. Los primers utilizados para detectar las diferentes formas del vector están indicadas con flechas. La pareja de primers Δ U3Fw y PBS Rv se usó para detectar la cantidad total de productos de RT (genomas del vector en forma de DNA). Los primers q2LTR Fw y q2LTR Rv se utilizaron para evaluar las cantidades relativas de círculos 2LTR frente a los genomas de DNA de vector total.

La inserción de fragmentos largos, como IS2, en el 3'LTR de IDLVs puede incrementar la RH entre LTRs durante el procesamiento de los genomas lineales de IDLVs, favoreciendo la formación de círculos con un solo LTR (círculos 1-LTR), los cuales se han asociado a mayores niveles de expresión comparado con los círculos 2-LTR²³⁰. Este representa, por tanto, otro de los mecanismos que podría explicar la superior actividad transcripcional de SE-IS2-IDLVs. Para contrastar esta posibilidad, se generó un IDLV control (SE-1.2Kb) albergando en el 3'LTR un fragmento irrelevante de DNA, equivalente a IS2 en longitud (1.2Kb) (Figura 21, a), y se comparó el efecto de la inclusión de ambos en IDLVs. Como ocurrió previamente en el caso de IS2, la inserción del fragmento de 1.2Kb no afectó al título de los vectores (número de partículas en el sobrenadante viral), estimado mediante RT-qPCR (Figura 21, b), pero redujo del orden de 4-5 veces la cantidad de productos de RT en células 293T (Figura 21, c).

La cantidad de círculos con 2-LTR se determinó mediante qPCR con la pareja de primers q2LTR Fw / q2LTR Rv, en células 293T 72h post-transducción. Del mismo modo, con la pareja de primers Δ U3 Fw / PBS Rv se determinó la cantidad de episomas totales. Tal y como se esperaba, el incremento en la longitud del LTR provocó una reducción en la cantidad de círculos con 2-LTR relativa a la cantidad de episomas totales, de forma similar en SE-IS2-IDLVs y SE-1.2Kb-IDLVs (Figura 21, e). Estos

resultados sugieren un incremento en círculos con 1-LTR en ambos casos. No obstante, al contrario de lo que sucede con SE-IS2-IDLV, la expresión de SE-1.2Kb-IDLV fue 2.5 veces inferior a la de SE-IDLV (Figura 21), dato que se correlacionaba con la reducción observada en el número de episomas (Figura 21, e). Estos resultados indican que la reducción en la proporción de círculos con 2-LTR (y, por tanto, el aumento relativo de las moléculas con 1-LTR), no es el principal mecanismo involucrado en el incremento de la actividad transcripcional de los episomas SE-IS2, dado que células 293T transducidas con SE-1.2Kb-IDLV mostraron una reducción similar en círculos con 2-LTR, sin mostrar un efecto positivo sobre la eficacia de transcripción.

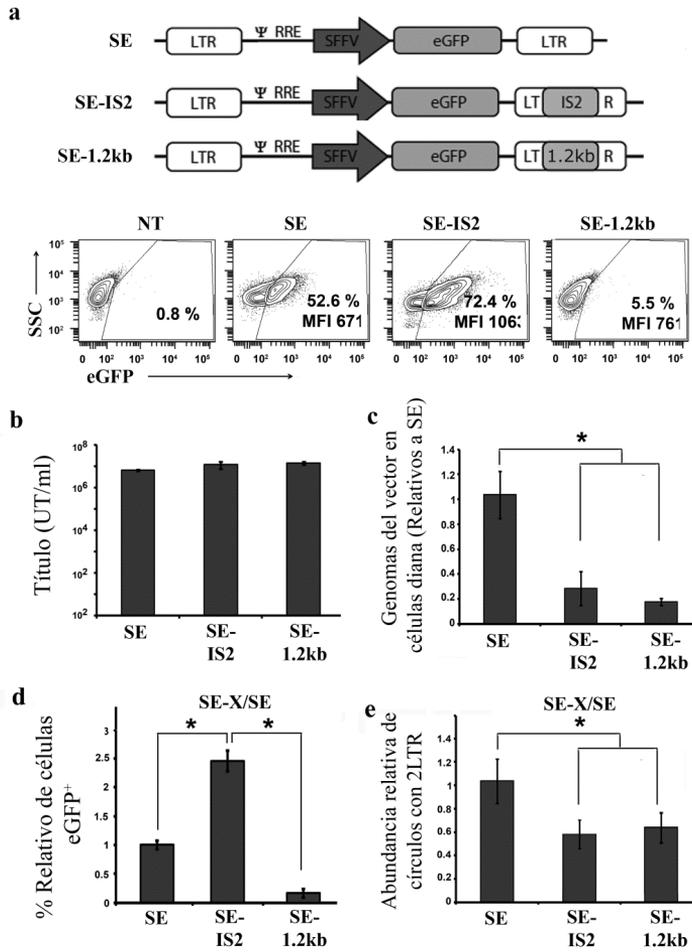


Figura 21. La inserción en IDLVs del elemento IS2, o de un fragmento control de 1.2 Kb, reduce el número total de episomas y la generación de círculos con 2-LTR, menos expresores. a) Arriba: representación esquemática de los vectores SE, SE-IS2 y SE-1.2Kb. Abajo: Plots representativos mostrando los perfiles de expresión de células 293T transducidas con los diferentes IDLVs, con un MOI de 0.7. Los porcentajes (%) y los niveles de expresión (MFI) de eGFP se muestran en cada plot. b) Gráficas mostrando títulos virales de los diferentes IDLVs (partículas virales presentes en el sobrenadante viral), determinados mediante RT-qPCR, usando el kit comercial para titulación de LV de ABM. c) Cantidad relativa de genomas del vector en células 293T transducidas con los diferentes IDLVs, determinada 72h post-transducción, y normalizada respecto a los niveles observados en células transducidas con SE-IDLVs. Los análisis se realizaron considerando un mismo número de células para todos los puntos, estimado mediante el número de copias del gen de la β -albúmina. d) Gráficas mostrando el porcentaje de células eGFP⁺ en 293T transducidas con los diferentes IDLVs, normalizado con respecto a los niveles observados en células transducidas con SE-IDLVs. e) Gráficas mostrando la abundancia relativa de círculos con 2-LTR con respecto al total de genomas del vector en células 293T, 72h post-transducción, normalizando frente a los niveles observados en 293T transducidas con SE-IDLVs. Los análisis se realizaron considerando un mismo número de células para todos los puntos, estimado mediante el número de copias del gen de la β -albúmina. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

1.13.3.6 La inclusión del elemento IS2 mantiene su efecto positivo en IDLVs que expresan el transgén a través del promotor fisiológico de WASP

Finalmente, analizamos si el elemento IS2 tenía un efecto similar cuando se insertaba en un esqueleto lentiviral diferente, que expresaba el transgén bajo un promotor fisiológico. Utilizamos un esqueleto LV desarrollado previamente por nuestro grupo (AWE)²⁰⁷, que expresa eGFP a través del promotor del gen *WAS* (gen codificante de la proteína del síndrome de Wiskott-Aldrich), solo en células hematopoyéticas²⁰⁶, para la generación de IS2-AWE (Figura 22, a). Dado que el promotor *WAS* es específico de linaje hematopoyético, probamos el comportamiento de AWE-IS2-IDLV frente al AWE-IDLV en células Jurkat (una línea inmortalizada de linfocitos T humanos). Como se puede observar en la Figura 22, b, los resultados fueron similares a los obtenidos previamente con los SE-IDLV. De hecho, aunque las células transducidas con AWE-IS2-IDLV tuvieron un número de episomas 3–4 veces más bajo en las células diana, en

comparación con las transducidas con AWE-IDLV (Figura 22, c), expresaron niveles similares o ligeramente más altos de eGFP (Figura 22, b).

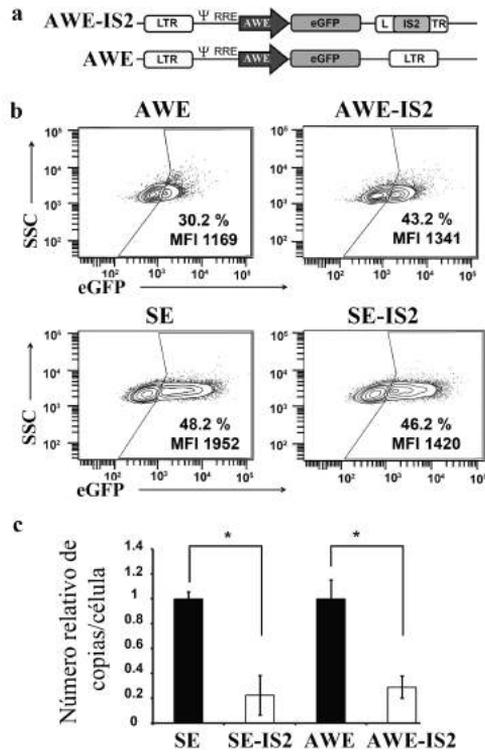


Figura 22. Rendimiento de IS2-IDLVs expresando el transgén a través del promotor humano de WASP. a) Esquema de los distintos IDLVs conteniendo el promotor de WASP. b) Plots representativos mostrando los perfiles de expresión de eGFP en células Jurkat transducidas con IDLVs AWE y AWE-IS2 (Plots superiores) y con SE y SE-IS2 (Plots inferiores), con un MOI de 0.5. Los porcentajes (%) de células eGFP+ y los niveles de expresión (MFI) se muestran en cada plot. c) Gráficas mostrando el número relativo de genomas del vector por célula en Jurkat transducidas con los distintos IDLVs, 72h post-transducción. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

1.13.3.7 El efecto final de la inclusión de IS2 en IDLVs depende de la célula diana

Tras estudiar el comportamiento de SE-IS2-IDLV en las líneas celulares K562 y 293T, se pasó a estudiar si la presencia del elemento IS2, favorece la expresión de IDLVs en nuestras células diana para TG de WAS (HSCs y células T), así como en otras células de interés en TG y/o en investigación básica, como lo son células progenitoras neuronales humanas (NPCs) y neuronas diferenciadas (NCs), células troncales pluripotentes inducidas (iPSCs), fibroblastos de piel humana (HSF) y células de mucosa oral humana (HOM).

El comportamiento de IS2 en IDLV se analizó en nuestras células diana (HSCs y células T), transduciendo estas células con SE- y SE-IS2-IDLVs (con un MOI de 5, en el caso de células T, y un MOI de 10 en HSCs), y analizando los niveles de expresión de eGFP mediante citometría de flujo, 72h post-transducción. Los efectos positivos en el patrón de expresión de IS2-IDLVs observados en las líneas celulares K562 y 293T no se reprodujeron en las células T, ni en las distintas subpoblaciones de células T analizadas (MOI 5) (Figura 23). Tampoco observamos diferencias en las HSCs, donde el efecto neto fue incluso negativo (Figura 24).

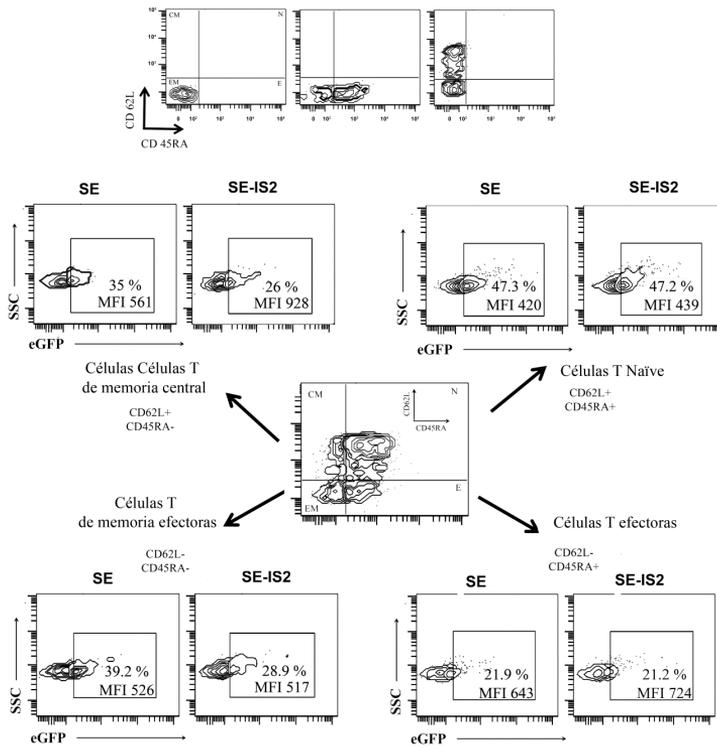


Figura 23. Comportamiento de IS2-IDLVs en varias subpoblaciones de células T. Los plots superiores muestran control de isotipo (izquierda), tinción simple con anti-CD45RA-PE (centro), y tinción simple con anti-CD62L-PE-Cy7 (derecha). Células T se transdujeron con IDLVs SE y SE-IS2 con un MOI de 5, y la expresión de eGFP se analizó 72h post-transducción. Las células transducidas se tiñeron con anti-CD45RA-PE y anti-CD62L-PE-Cy7, y la expresión de eGFP se analizó en cada una de las distintas subpoblaciones: células T de memoria efectoras (CD62L-CD45RA-), células T efectoras (CD62L-CD45RA+), células T de memoria central (CD62L+CD45RA-), y células T Naïve (CD62L+CD45RA+).

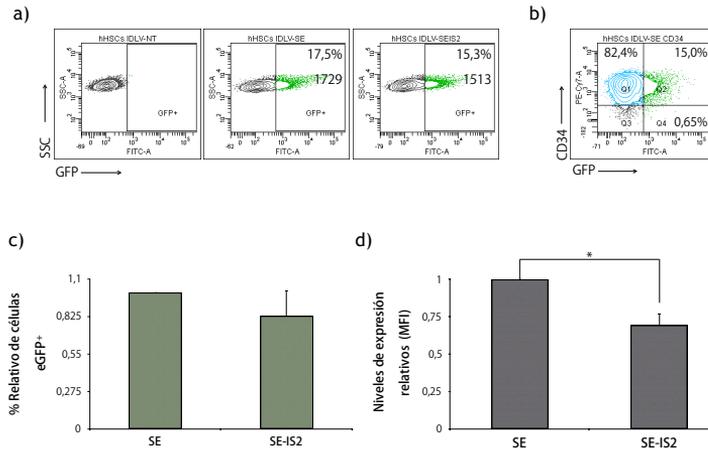


Figura 24. Rendimiento de IS2-IDLVs en HSCs. a) Plots representativos mostrando los perfiles de expresión de eGFP en HSCs procedentes de sangre periférica movilizada (PMB), transducidas con SE (centro) y SE-IS2 (derecha) IDLVs, con un MOI de 10, y no transducidas (izquierda), como control. La expresión de eGFP se analizó 72h post-transducción. Los porcentajes (%) y los niveles de expresión (MFI) de eGFP se muestran en cada plot. b) Plot representativo mostrando fenotipo CD34⁺ de las HSCs transducidas con IDLV. c) Gráfica mostrando el porcentaje de células eGFP⁺ en HSCs transducidas con SE-IS2-IDLVs, normalizado frente a los niveles observados en células transducidas con SE-IDLVs. d) Gráfica mostrando los niveles de expresión relativos (MFI) de HSCs transducidas con SE-IS2-IDLVs, normalizados frente a los niveles de expresión de SE-IDLVs. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

Tras los datos decepcionantes del efecto de IS2-IDLVs en HSCs y células T, nos propusimos estudiar si los IS2-IDLVs podrían ser de aplicación en otras células especialmente relevantes en medicina regenerativa. Comenzamos analizando fibroblastos humanos derivados de piel y de mucosa oral. En ambos casos, usando un MOI de 10 partículas virales por célula para las transducciones, tampoco observamos un efecto final positivo de SE-IS2-IDLVs sobre los niveles de expresión (Figura 25).

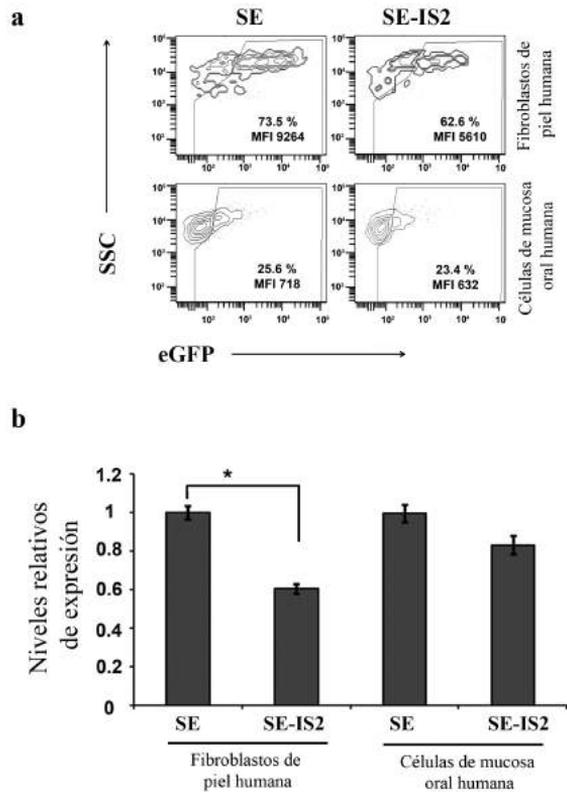


Figura 25. Rendimiento de IS2-IDLVs en fibroblastos humanos derivados de piel, y de mucosa oral. **a)** Plots representativos mostrando los perfiles de expresión de fibroblastos humanos procedentes de piel (plots superiores) y de mucosa oral (plots inferiores), transducidos con SE (izquierda) y SE-IS2 (derecha) IDLVs, con un MOI de 10. La expresión de eGFP se analizó 72h post-transducción. Los porcentajes (%) y los niveles de expresión (MFI) de eGFP se muestran en cada plot. **b)** Gráficas mostrando los niveles relativos de expresión de SE-IS2-IDLVs en ambos tipos celulares, normalizados frente a los niveles de expresión de SE-IDLVs en la misma línea celular con el mismo MOI. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

Estos resultados nos indican que el efecto neto del IS2 en los niveles de expresión de los IDLVs en Células T, HSCs y fibroblastos humanos es negativo y, por lo tanto, su utilización en estas células no es pertinente.

1.13.3.7.1 La inclusión de IS2 en IDLV mejora los niveles de expresión en NPCs y neuronas diferenciadas

NPCs y neuronas diferenciadas se ven afectadas en trastornos mentales, como esquizofrenia y trastornos del espectro autista, enfermedades neurodegenerativas, como Parkinson y Alzheimer, y en ciertos tumores cerebrales²³¹. Son, por tanto, células de especial interés para TG mediante EG y/o para el desarrollo de modelos.

Para analizar el comportamiento de IS2-IDLVs en NPCs, se transdujeron estas células con los IDLVs SE, SE-IS2, SEWP y SEWP-IS2 con un MOI de 3, y se determinaron los niveles de expresión de eGFP mediante citometría de flujo, 72h post-transducción. Estos niveles de expresión se normalizaron con respecto a los niveles de expresión observados en NPCs transducidas con SE-IDLVs. Adicionalmente, se llevó a cabo la diferenciación neuronal *in vitro* a partir de NPCs (ver material y métodos), para reproducir estos análisis en neuronas diferenciadas. Tras 28 días de diferenciación, las neuronas diferenciadas se fenotiparon mediante inmunofluorescencia, con el anticuerpo monoclonal Tuj1 frente a β III tubulina, marcador característico neuronal²³²⁻²³⁵ (Figura 26).

Diferenciación neuronal a partir de
Células progenitoras neuronales humanas derivadas de H9

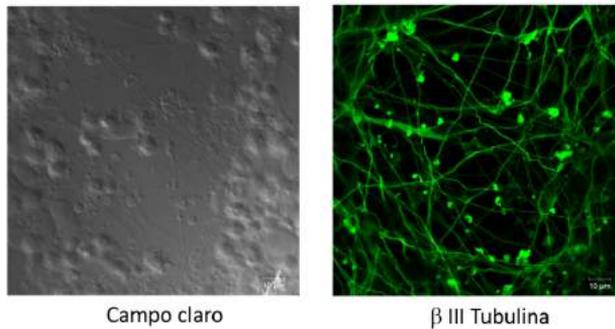


Figura 26. Imagen de neuronas diferenciadas a partir de NPCs. NPCs se cultivaron durante 28 días en medio de diferenciación neuronal, a 37°C en una atmósfera del 5% de CO₂. Transcurrido este tiempo, las células se fijaron y permeabilizaron, y se tiñeron con el anticuerpo monoclonal de ratón Tuj1 frente a β III tubulina, y un anticuerpo secundario de cabra anti-IgG de ratón conjugado con Alexa Flúor 488. Las imágenes de microscopía confocal se tomaron con el microscopio de barrido láser confocal Zeiss LSM 710 acoplado a un microscopio vertical Axiomager A1.

La inclusión del elemento IS2 en IDLVs reflejó un marcado incremento en la expresión de eGFP en NPCs, con independencia de que el esqueleto lentiviral incluyese o no el regulador postranscripcional WPRE (Figura 27, c). Sin embargo, en neuronas diferenciadas la inclusión de IS2 en IDLV solo mostró un incremento en los niveles de expresión de eGFP en ausencia de WPRE (MFI SE=797 *versus* SE-IS2=2178) (Figura 27, d). Este efecto positivo era enmascarado por la presencia de WPRE.

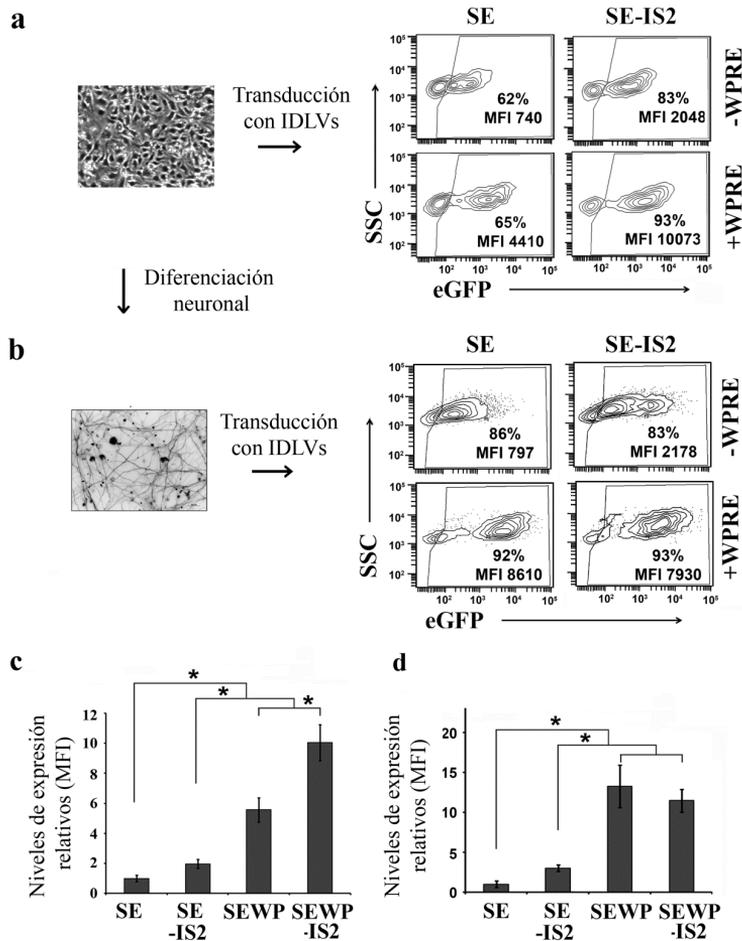


Figura 27. El elemento IS2 mejora los niveles de expresión de eGFP en células progenitoras neuronales humanas (NPCs) y en neuronas diferenciadas, transducidas con IDLVs. a) Imagen por microscopía de transmisión (izquierda) y plots (derecha) mostrando perfiles de expresión de eGFP en NPCs transducidas con SE, SE-IS2, SEWP y SEWP-IS2 IDLVs, con un MOI de 3, analizados 72h post-transducción. En cada plot se muestra el porcentaje (%) de eGFP y los niveles de expresión (MFI). b)

Imagen por microscopía de transmisión (izquierda) y plots (derecha) mostrando perfiles de expresión de eGFP de neuronas diferenciadas (positivas para β III tubulina) transducidas con SE, SE-IS2, SEWP y SEWP-IS2 IDLVs, con un MOI de 3, analizados 72h post-transducción. En cada plot se muestra el porcentaje (%) de eGFP y los niveles de expresión (MFI). c) Gráficas mostrando los niveles de expresión de eGFP relativos en NPCs transducidas con los distintos IDLVs, y normalizados con respecto a los niveles observados en NPCs transducidas con SE-IDLVs. d) Gráficas mostrando los niveles de expresión de eGFP relativos en neuronas diferenciadas transducidas con los distintos IDLVs, y normalizados con respecto a los niveles observados en NPCs transducidas con SE-IDLVs. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

1.13.3.7.2 IS2-IDLVs mejoran la eficiencia de transducción y los niveles de expresión en iPSCs

Otro de los tipos celulares sobre los que se quiso analizar el posible efecto de la incorporación de IS2 en IDLVs fue en células troncales pluripotentes inducidas (iPSCs), por su relevancia para entender los mecanismos moleculares que regulan la pluripotencia, su utilidad para generar modelos de enfermedad, y sus oportunidades terapéuticas. Por tratarse de células complicadas de modificar²³⁶, para estudiar el efecto de IS2-IDLVs en iPSCs se usó un MOI de 10 en las transducciones con los distintos IDLVs. Transcurridas 72h desde la transducción, se comprobó que la incorporación de IS2 en el esqueleto de IDLVs causó en iPSCs un incremento significativo tanto del porcentaje relativo de células expresando eGFP (Figura 28, b, gráfica superior), como de los niveles de expresión (Figura 28, b, gráfica inferior). El porcentaje relativo de células eGFP⁺ se vio incrementado tanto en ausencia como en presencia (Figura 28, b, gráfica superior) de WPRE.

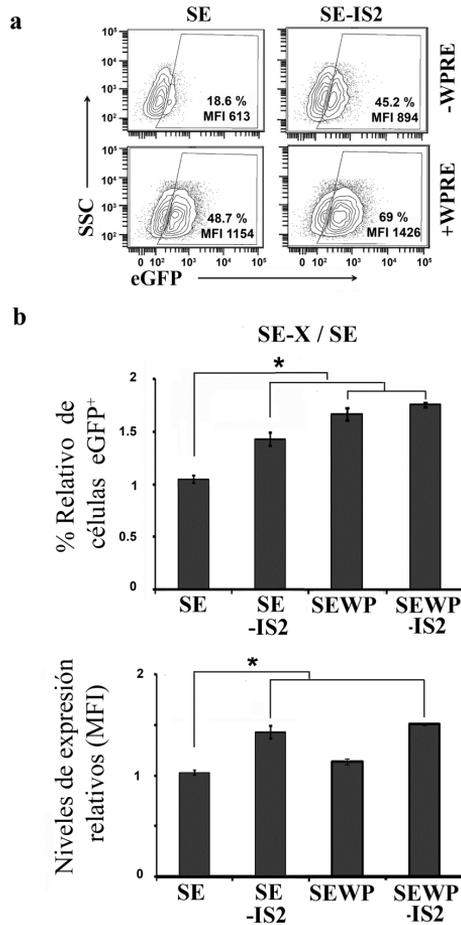


Figura 28. IS2 mejora los niveles de expresión de IDLVs en iPSCs. a) Plots representativos mostrando los perfiles de expresión de eGFP en iPSCs transducidas con SE, SE-IS2, SEWP, y SEWP-IS2 -IDLVs, con un MOI de 10. El porcentaje (%) de eGFP y los niveles de expresión (MFI) se indican en cada plot. b) Gráficas mostrando los porcentajes relativos de células eGFP+ (gráfica superior) y los niveles de expresión relativos (gráfica inferior) en iPSCs transducidas con los diferentes IDLVs, y normalizadas en base a los niveles obtenidos en iPSCs transducidas con SE-IDLVs. Los datos representan la media \pm el error estándar de al menos 4 experimentos independientes (* $p < 0.05$).

1.13.3.8 Inclusión de IS2 en IDLVs para entrega de Cas9 en HSCs y K562.

A pesar de que la inclusión de IS2 en IDLVs no mostró un efecto claro sobre los niveles de expresión de eGFP en nuestras células diana, decidimos incluir este elemento en el 3'LTR del plásmido lentiviral Lgl₉Cas9, que permite la entrega del sistema CRISPR/Cas9 gl₉ (frente al primer intrón del gen *WAS* (Figura 29, a), a las células diana en forma de IDLV, para ver si sobre la expresión de Cas9 se producía un efecto neto positivo.

Una vez conseguido el plásmido lentiviral Lgl₉Cas9_IS2, originado a partir de Lgl₉Cas9, comparamos ambos IDLVs en nuestras células modelo (K562), y en nuestras células diana (HSCs). Una semana después de las transducciones, se analizó en ambos tipos celulares la eficiencia de corte en el locus *WAS* de Lgl₉Cas9 y Lgl₉Cas9_IS2. En K562, ambos IDLV interrumpían el primer intrón de *WAS* de forma eficiente, sin encontrar diferencias sustanciales entre ambos IDLVs, en cuanto a eficiencia de corte. Estos datos no mimetizan los datos obtenidos con eGFP, por lo que es posible que al incluir Cas9 (de mayor tamaño que eGFP) se esté afectando la viabilidad de los IS2-IDLVs (ya de por sí de elevado tamaño). Por último, tampoco pudimos observar, como era de esperar por los datos en K562, que ningún IDLV-Cas9 consiguiera generar indels en HSCs (Figura 29, b).

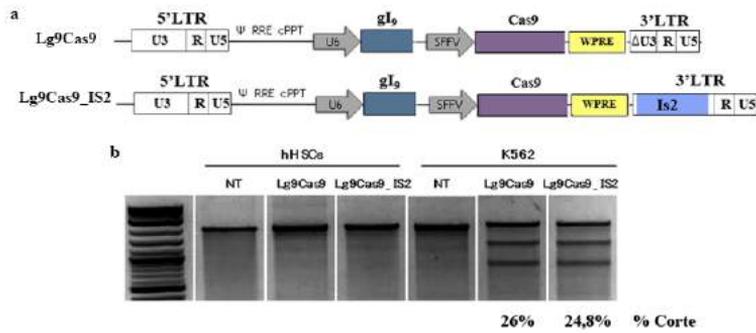


Figura 29. Efecto del elemento IS2 en Cas9_IDLVs, en HSCs y K562. a) Esquema representativo de los plásmidos lentivirales Lg9SCas9 y Lg9SCas9_IS2. b) Eficiencia de corte de los IDLV Lg9SCas9 y Lg9SCas9_IS2 en nuestras células diana (HSCs), y en el modelo celular K562, determinada mediante el ensayo de la t7.

1.14 OBJETIVO 2: DESARROLLO DE MODELOS PARA ESTUDIAR EFICIENCIA Y SEGURIDAD DE LA REPARACIÓN GÉNICA MEDIANTE CRISPR/Cas9

Existen diferentes estrategias para reparar la expresión de los genes mutados, sin embargo, la estrategia más precisa (y por tanto la más deseada) requiere de un DNA donador que introduzca en el sitio deseado la secuencia correcta. Para trasladar estas tecnologías a pacientes es necesario minimizar al máximo tanto los "*off targets*" de las nucleasas específicas como las integraciones al azar del DNA donador utilizado. Esto es así porque ambas inespecificidades podrían dar lugar a efectos indeseados en las células diana. Por lo tanto, otro objetivo de esta tesis fue analizar en detalle la influencia del diseño del DNA donador y/o los métodos de entrega de los diferentes elementos del sistema para lograr una mayor eficiencia y especificidad de edición. Para ello decidimos en primer lugar desarrollar dos modelos celulares diferentes que nos permitan estudiar la eficacia y la seguridad de reparación mediante recombinación homóloga. Estos modelos fueron después utilizados para determinar que sistemas nos ofrecen una mayor eficacia y una mejor especificidad de edición genómica.

1.14.1 Generación de la línea reportera K562 SE- δ GF-I₁-P

Nuestro primer modelo celular se basa en el plásmido lentiviral SEWP (Figura 12), donde se mutó la secuencia codificante para eGFP, mediante la eliminación del primer ATG y la introducción de un fragmento de 387 pb del primer intrón (I₁) del locus *WAS*. En esta secuencia se encuentra la diana para el sistema sistema CRISPR/Cas9 gI9. La obtención de esta construcción, denominada SE- δ GF-I₁-P (Figura 30, a), se detalla en material y métodos.

La línea celular K562- δ GF-I₁-P se generó mediante la transducción de K562 con un MOI aproximado de 20 del LV SE- δ GF-I₁-P. La línea celular K562 se escogió por ser una línea de linaje hematopoyético, cuya accesibilidad, fácil cultivo, y transfectabilidad, la convierten en idónea para la generación de modelos para EG de *WAS*. La línea obtenida, K562- δ GF-I₁-P, presentaba integrado en su genoma una media de 13 copias del cassette SE- δ GF-I₁-P (estimadas mediante qPCR con la pareja de primers Δ U3Fw / PBS Rv, ver M&M) y carecía por completo de expresión de eGFP (Figura 30, b). El alto MOI utilizado para conseguir >10 copias por células tenía la finalidad de conseguir un número alto de integraciones del cassette SE- δ GF-I₁-P que dotasen de una mayor sensibilidad para la RH en la línea reportera. De esta manera, la línea K562- δ GF-I₁-P permite analizar la eficacia de la edición genómica utilizando sistemas específicos del locus *WAS*, pero analizando simplemente el rescate de la expresión de eGFP. Además, diseñando adecuadamente el DNA donador, se puede estudiar simultáneamente la eficacia y seguridad (especificidad) de diferentes sistemas simplemente analizando el patrón de expresión de proteínas fluorescentes (ver más adelante).

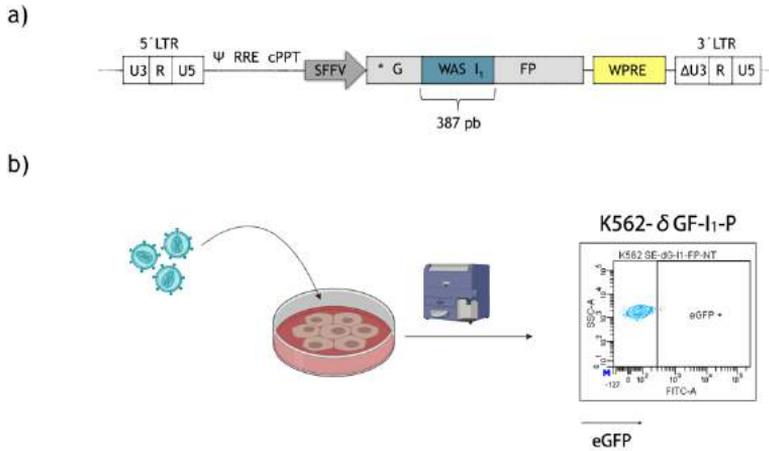


Figura 30. Obtención de la línea celular K562- δ GF-I₁-P, reportera de RH. a) Esquema plásmido lentiviral SE- δ GF-I₁-P conteniendo la secuencia de eGFP mutada e interrumpida por un fragmento de 387 pb del locus *WAS*, que incluye la diana para el sistema CRISPR/Cas9 gI₉. **b)** Esquema representativo de la estrategia seguida para la generación de esta línea. Células K562wt se transdujeron con SE- δ GF-I₁-P LV, con un MOI de 20, y posteriormente se comprobó la ausencia de expresión de eGFP en estas células, conteniendo varias copias del cassette SE- δ GF-I₁-P integradas en su genoma.

1.14.1.1 Análisis de la eficacia y seguridad de EG en el modelo K562- δ GF-I₁-P

Una vez generado la línea celular reportera K562- δ GF-I₁-P, diseñamos un donador que permitiese determinar la eficacia y especificidad de corrección génica. Este donador debería por tanto ser capaz de indicar cuándo se producía la recombinación homologa de forma “correcta” y mostrar el porcentaje de inserciones “ilegítimas” del donador en sitios distintos a la diana. Un donador que contenga el cassette SFFV_eGFP

(SE) seguido de otro cassette expresando una proteína fluorescente alternativa, cumpliría estos requisitos. El cassette SE contiene los brazos de homología y la secuencia correcta de eGFP que nos indicará la presencia de recombinación homóloga. El otro cassette de expresión solo se integraría al azar o mediante procesos diferentes a la RH, y nos indicaría por tanto integraciones ilegítimas del DNA donador (Figura 31).

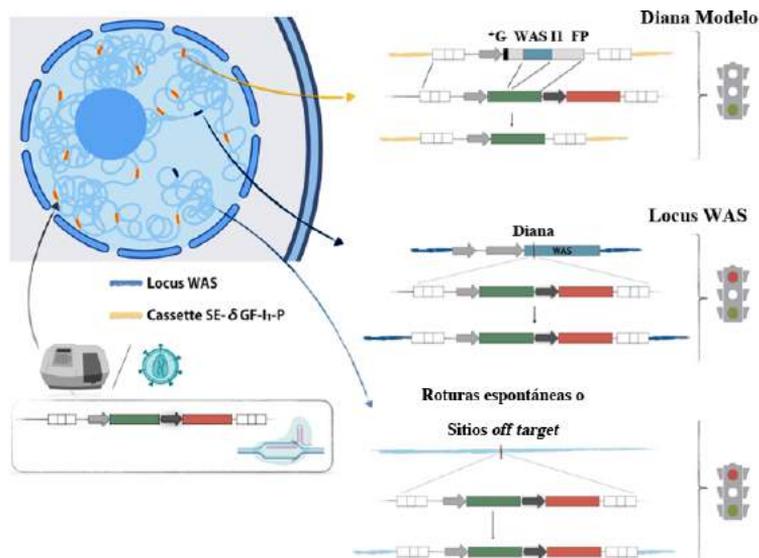


Figura 31. Ilustración esquematizando, de forma simplificada, el tipo de inserciones del donador que pueden producirse en el genoma de la línea celular K562- δ GF-I₁-P. En la parte superior izquierda aparece representado el núcleo de células K562- δ GF-I₁-P, con un promedio de 13 integraciones del cassette SE- δ GF-I₁-P en su genoma (marcadas en amarillo), y dos copias del locus WAS (marcadas en azul). A la derecha se esquematizan los tres tipos de integraciones del donador que podrían producirse: 1- integración del donador mediante RH en la diana (cassette SE- δ GF-I₁-P), que supondría una restauración de la expresión de eGFP, 2- integración del donador en el locus WAS endógeno, donde se localiza la diana original del sistema CRISPR/Cas9 gI9, y 3- integración del donador en otros sitios ilegítimos del genoma en los que el sistema CRISPR/Cas9 gI9 esté

produciendo cortes (sitios *off target*), o en los que se hayan producido roturas espontáneas del DNA en presencia del donador.

Con esta idea se generó la construcción SEED (Figura 32) en un contexto lentiviral para poder hacer la entrega del DNA donador tanto como plásmido, como mediante IDLVs. En el donador SEED, la secuencia codificante de dsRed se encuentra bajo el promotor del factor de elongación 1 α humano (EF1 α), aguas abajo y en la misma dirección que el cassette SE. De esta forma, aquellas células en que el donador se hubiese insertado en sitios distintos a la diana expresarían ambas proteínas (eGFP y dsRED), mientras que aquellas en que únicamente se hubiese producido RH serían solo eGFP⁺, una vez extinguida la expresión episomal del donador. Un potencial problema del donador SEED son las 236pb de homología en el extremo 3' que podrían utilizarse para RH y que darían como resultado células eGFP⁺ dsRED⁺, sobrestimándose entonces las inserciones fuera de sitio. Sin embargo, como el cassette EF1 α –dsRed tiene más de 2kb, hipotetizamos que estos sucesos serían mínimos, teniendo en cuenta la lejanía de esta región al punto de corte de nuestro sistema CRISPR/Cas9.

En conjunto, este modelo celular nos dará una comparativa no-sesgada de los sitios de integración del DNA donador, dado que puede identificar sitios de integración tanto en sitios *off target* del sistema CRISPR, como en sitios donde el donador se haya integrado por cualquier otro proceso. Este modelo fue en primer lugar utilizado para comparar la eficacia y especificidad de EG de diferentes sistemas de entrega del sistema CRISPR/Cas9 específico para *WAS*. A continuación, se procedió a analizar la influencia del diseño del DNA donador, así como de los métodos de entrega de éste.

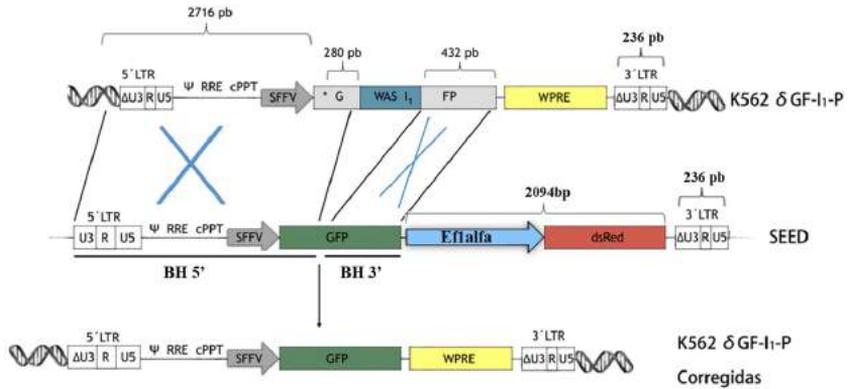


Figura 32. Restauración de la secuencia de eGFP en la línea celular K562-δGF-I1-P. En el diagrama superior de la figura se representa el genoma de la línea celular K562-δGF-I1-P, en aquellas posiciones en que se encuentra integrada la construcción SE-δGF-I1-P, que contiene la diana del I₁ de *WAS* para el sistema CRISPR/Cas9 gI9. El diagrama central muestra el plásmido donador SEED, compuesto por el cassette de expresión SE, para promover la RH y rescatar la expresión de eGFP en K562-δGF-I1-P, y por el cassette EF1α-dsRed, para identificar integraciones del donador que se produzcan fuera de la diana. El diagrama inferior muestra el locus resultante después de que se produzca la reparación de eGFP por RH, utilizando los brazos de homología indicados (BH 5' y BH 3').

1.14.1.1.1 Comparativa de la eficacia y especificidad de la EG de diferentes sistemas de entrega del sistema CRISPR/Cas9 utilizando el donador SEED

Sobre la línea celular K562-δGF-I1-P, utilizando el plásmido SEED como donador, se compararon eficiencia y seguridad de corrección génica de eGFP conseguidas con el sistema CRISPR/Cas9 gI9, empleando tres métodos distintos para vehicular este sistema hasta las células diana: 1) Nucleofección de DNA donador (plásmido SEED) y Lgl9Cas9 (plásmido), 2) Nucleofección de DNA donador (plásmido SEED) y de ribonucleoproteína (RNP gI9) y 3) Nucleofección de DNA donador (plásmido SEED) y transducción con IDLVs Lgl9Cas9. Como control negativo, se nucleofectaron células

K562- δ GF-I₁-P con el plásmido SEED, en ausencia del sistema CRISPR/Cas9 g₉. La expresión de eGFP y dsRed comenzó a medirse, mediante citometría de flujo, 48 horas después de proporcionar a las células K562- δ GF-I₁-P, donador y nucleasa, y se analizó de forma periódica, hasta su estabilización, una vez diluida la señal episomal transitoria del donador (Figura 33). La eficiencia relativa de RH se determinó relativizando el porcentaje de células eGFP⁺ dsRed⁻ a tiempo final respecto a la eficacia de entrada del donador en las células diana (definida como el porcentaje de células eGFP⁺ dsRed⁺ a las 48h). Al valor resultante se sustrajo el obtenido en ausencia del sistema CRISPR/Cas9 g₉.

La expresión de SEED desapareció prácticamente por completo en células a las que no se administró CRISPR/Cas9 g₉, ratificando que ambos elementos son necesarios para que se produzca la integración del donador en el genoma de la célula diana de forma eficiente. Por otro lado, al suministrar donador y nucleasa, la eficacia de RH (% de células eGFP⁺ dsRed⁻) fue 100 – 300 veces superior al de inserciones inespecíficas del donador (% de células eGFP⁺ dsRed⁺), dando una especificidad de RH superior al 99% en el peor de los casos (Figura 33, d). Estos números no reflejaban estudios anteriores del grupo, donde la especificidad de la recombinación homóloga nunca era mayor del 90%²³⁷. Creemos que esto podría ser debido al elevado número de sitios diana (13) encontrados en este modelo, favoreciendo los procesos de RH y reduciendo así las integraciones inespecíficas.

Cuando comparamos los diferentes métodos empleados para la entrega del sistema CRISPR/Cas9 g₉ en términos de eficiencia de RH, se pudo observar que las RNP eran ligeramente superiores, aunque las diferencias no fueron significativas (Figura 33, b). Lo mismo ocurriría al medir el % de inserciones ilegítimas (Figura 33, c). Sin embargo, estas tendencias a favor de las RNP desaparecen cuando se analiza la especificidad

(relativizando el % las inserciones ilegítimas entre el % de RH), indicando que todos los sistemas de entrega ofrecen una especificidad similar (Figura 33, d).

Resultados

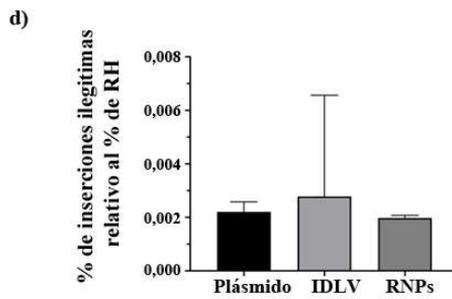
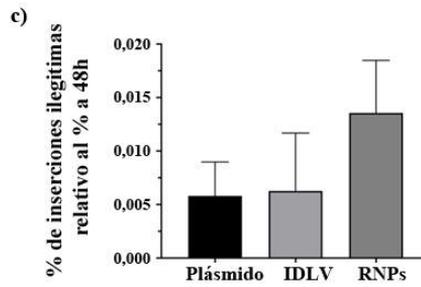
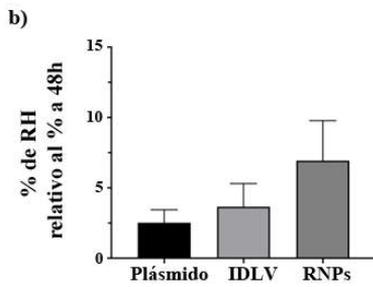
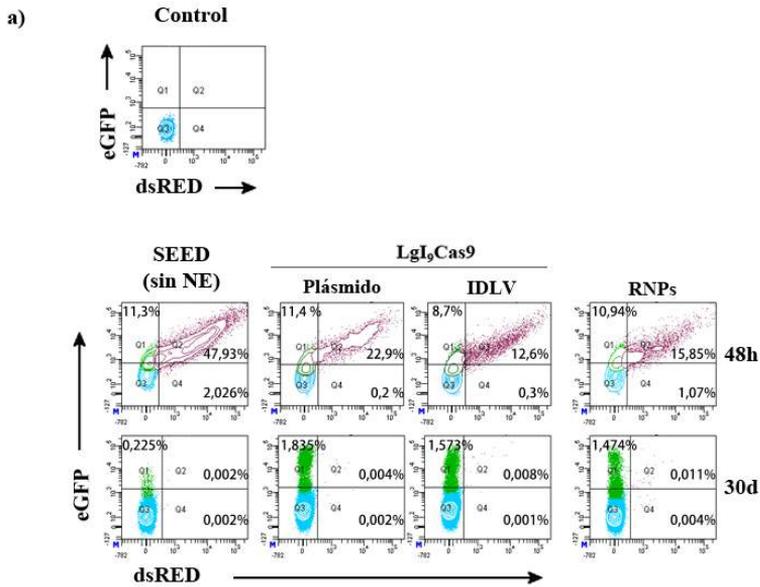


Figura 33. Rescate de eGFP en línea celular reportera K562 SE- δ GF-I₁-P, mediante distintos métodos de entrega del sistema CRISPR/Cas9 g₉. a) Plots de citometría representativos mostrando seguimiento en el tiempo de la expresión, tanto de eGFP, como de dsRed. El plot superior muestra las células del modelo sin transfectar. Los dos plots inferiores-izquierda muestran las células K562 SE- δ GF-I₁-P nucleofectadas únicamente con el donador SEED. El sistema CRISPR/ Cas9 g₉ se entregó a las células mediante tres métodos distintos: 1- como plásmido y mediante co-nucleofección con el donador (plots inferiores 2do-izquierda), 2- como IDLV, transduciendo las células inmediatamente después de nucleofectarlas con el donador (plots inferiores 2do-derecha), y 3- como RNP, mediante co-nucleofectando junto con el donador (plots inferiores-derecha). Se hizo una lectura inicial de la expresión de eGFP y dsRed transcurridas 48h (plots superiores), y otra a día 30 (plots inferiores). b) Gráfico mostrando la eficacia relativa de RH en el modelo K562 SE- δ GF-I₁-P con cada uno de los sistemas de entrega. La eficacia relativa de RH se calculó relativizando el % de células eGFP⁺ a d30 con respecto al % de células eGFP⁺dsRED⁺ a 48h, y extrayendo el valor obtenido para el control negativo (células nucleofectadas sólo con el donador SEED). c) Gráficas mostrando la frecuencia relativa de inserciones ilegítimas (sin RH), calculada relativizando el % de células eGFP⁺dsRED⁺ a d30 con respecto al % de células eGFP⁺dsRED⁺ a 48h, y extrayendo el valor obtenido para el control negativo (células nucleofectadas sólo con el donador SEED). d) Gráfico mostrando la especificidad de los diferentes sistemas de entrega calculada en base a la relación entre la frecuencia relativa de integraciones ilegítimas y la eficiencia relativa de RH. Los datos representan la media \pm el error estándar de al menos 3 experimentos independientes. La comparación entre los distintos grupos se llevó a cabo mediante el método no paramétrico de Dunn (*p < 0.05).

1.14.1.1.2 Comparativa de la eficacia y especificidad de la EG de diferentes donadores y diferentes sistemas de entrega

A continuación, se procedió a investigar la influencia del diseño del DNA donador, así como el método de entrega en la eficacia y especificidad de la edición. Se generaron tres construcciones distintas, además del donador SEED, albergada cada una de ellas en un plásmido lentiviral, con la concepción inicial de poder entregar estos donadores a las células diana tanto en forma de plásmido, como en forma de IDLV.

Todos los donadores que se diseñaron contienen el cassette de expresión SFFV-eGFP, seguido de un cassette de expresión de la proteína roja fluorescente dsRed²³⁸, de forma que, al igual que con el donador SEED, aquellas células en que el donador se

hubiese insertado en sitios distintos a la diana expresarían ambas proteínas, mientras que aquellas en que sólo se hubiese producido RH serían eGFP⁺dsRed⁻, una vez extinguida la expresión episomal del donador. Los diferentes donadores diseñados son (Figura 34).

SECdR: En este donador la secuencia codificante de dsRed se encuentra bajo el promotor temprano de Citomegalovirus humano (CMV), aguas abajo del cassette SFFV-eGFP, y en antisentido con respecto a este. En esta construcción, el cassette CMV-dsRec (CdR) incluye la señal de poliadenilación (pA) del gen de la hormona de crecimiento bovina (bGH).

SECdR_IS2: Uno de los factores determinantes en la capacidad de unión de Cas9 a su diana es el estado de la cromatina, de forma que regiones heterocromáticas afectan negativamente a esta unión, mientras que un mayor nivel de apertura de la cromatina la favorece^{109,239-246}. Por otro lado, nosotros habíamos comprobado que la inclusión del elemento IS2 en IDLVs favorecía en células transducidas un reposicionamiento de los episomas hacia regiones de menor densidad cromatínica del núcleo²⁴⁷. En base a estas observaciones, pensamos que la inclusión de IS2 en un plásmido donador podría tener un efecto similar, dirigiendo el donador en el núcleo hacia regiones con menor densidad de la cromatina, donde la actividad de Cas9 será previsiblemente mayor. Por esta razón, se desarrolló el donador SECdR_IS2, mediante la incorporación del elemento IS2 en el 3'LTR del plásmido lentiviral SECdR.

SECdR_WI1R: Algunos autores han reportado eficacias de RH mayores al usar como donadores plásmidos que contienen el sitio diana de la nucleasa específica incluida en el DNA donador^{248,249}. Por ello, nos propusimos incluir la diana para nuestro sistema CRISPR/Cas9 g₉ aguas arriba del cassette SFFV-eGFP. Esto se consiguió clonando el

fragmento de 387pb del I₁ de *WAS* empleado para la generación del plásmido SE- δ GF-I₁-P aguas arriba del cassette SE en el donador SECdR, y en antisentido con respecto a este. Sin embargo en este caso, hay que tener en cuenta que esto, además de incrementar la frecuencia de recombinación homóloga en nuestra diana, incrementará la frecuencia de inserciones en el locus *WAS* endógeno, por microhomologías entre en sitio de corte en el donador y el locus *WAS* endógeno.

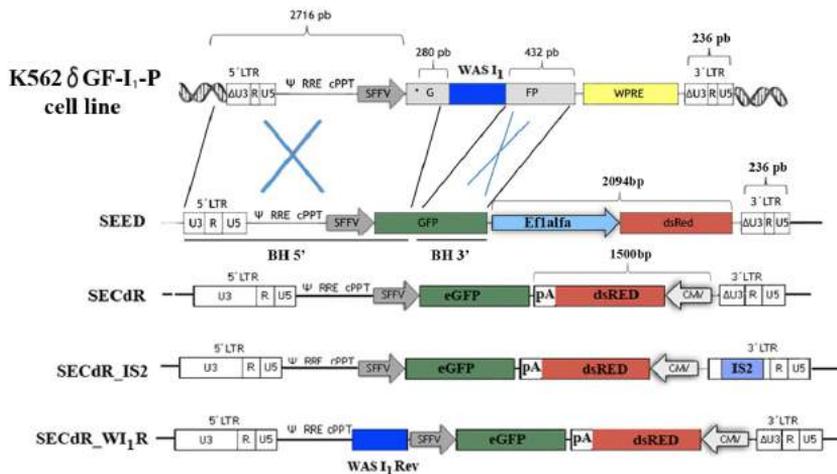


Figura 34. Diseño de donadores para restauración de eGFP en línea celular K562- δ G-I₁-FP. En el diagrama superior de la figura se representa el genoma de la línea celular K562- δ G-I₁-FP, en aquellas posiciones en que se encuentra integrado el cassette SE- δ GF-I₁-P, que contiene la diana para el sistema CRISPR/Cas9 g₉. En la parte inferior se representan los distintos plásmidos lentivirales empleados como donadores. El donador SEED está compuesto por los cassettes de expresión SFFV_eGFP (SE) y EF1 α _dsRed, ambos en la misma orientación. El donador SECdR, compuesto por el cassette SE, y el cassette CMV_dsRed_pA en orientación inversa. En base a SECdR se crearon dos donadores adicionales. Por un lado, el donador SECdR_IS2, conteniendo el elemento IS2 en el 3'LTR del plásmido lentiviral SECdR, y por otro, el donador SECdR_WI₁R, conteniendo el mismo fragmento del I₁ de *WAS* de 387pb que interrumpe la secuencia de eGFP en el modelo K562- δ GF-I₁-P, y que incluye la diana para el sistema CRISPR/Cas9 g₉.

Cada uno de los donadores diseñados (SEED, SECdR, SECdR_IS2 y SECdR_WI1R) (Figura 34) se entregó a células K562- δ GF-I1-P mediante nucleofección del plásmido sólo (como control), o en combinación con el sistema CRISPR/Cas9, entregado también como plásmido.

El seguimiento se hizo igual que experimentos anteriores, hasta la estabilización de los niveles de expresión de eGFP y dsRED, y se determinó la eficiencia relativa de RH como se ha explicado anteriormente. Los datos mostraron que la única modificación que mostró un incremento significativo de la eficacia de RH fue SECdR_WI1R. Este donador, que incluía la secuencia diana para CRISPR/Cas9, conseguía eficacias relativas de RH cercanas al 60% (y absolutas, cercanas al 20%) (Figura 35).

Hay que tener en cuenta que, al contener el donador SECdR_WI1R la diana para el sistema CRISPR/Cas9 gI9 aguas arriba del cassette de expresión SE, se podrían estar produciendo dos fenómenos: por un lado la generación de DSBs en el donador y en el locus diana favorecería la RH, incrementado la eficacia, y dando lugar a más células eGFP⁺dsRED⁺. Sin embargo, cortes en el locus *WAS* endógeno podrían dar lugar a inserciones ilegítimas mediante inserción mediada por aEJ utilizando regiones de microhomología (Figura 6), aunque estas se vean reducidas por la incorporación en reverso de la secuencia del intrón 1 de *WAS* en el donador. Es importante hacer notar que, a pesar de esto, el donador SECdR_WI1R ofrecía una especificidad relativa similar a los donadores SEED, SECdR y SECdRIS2 (Figura 35, d).

Por otro lado, otro factor que podría estar favoreciendo la mayor eficacia del donador WAS I₁ Rv_SECdR a la hora de rescatar eGFP en el modelo K562 δ GF-I1-P, podría ser la reducción en el tamaño del brazo de homología en 5' con respecto al resto de donadores (de 2716pb a 688pb), debida al corte que estaría produciendo el sistema CRISPR/Cas9 gI9 en su diana incluida en este donador.

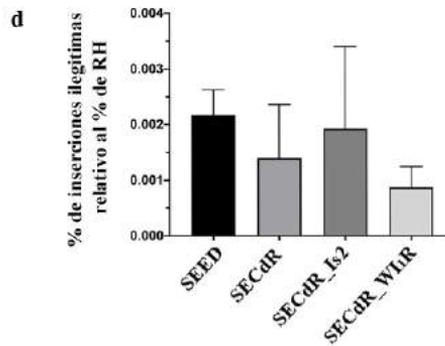
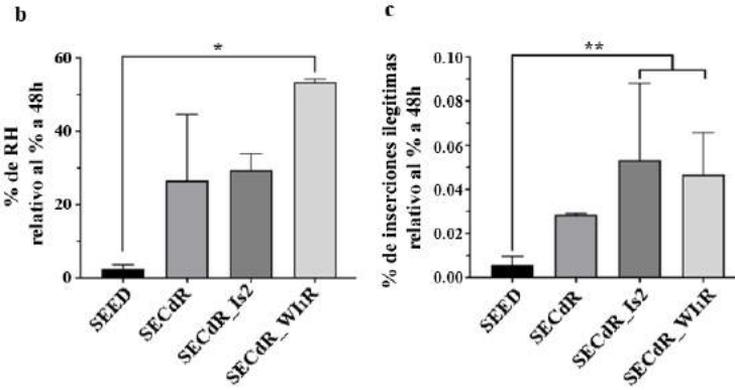
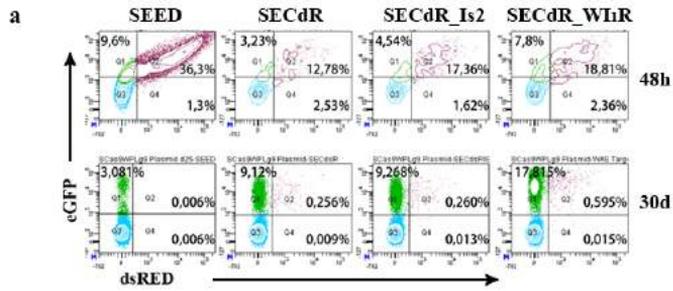


Figura 35. Eficacia y especificidad de la EG en K562- δ GF-I₁-P de diferentes donadores. a) Plots de citometría representativos mostrando seguimiento en el tiempo de la expresión, tanto de eGFP (FITC-A), como de dsRed (PE-A), en K562 SE- δ GF-I₁-P. Tanto los distintos donadores (SEED, SECdR, SECdR_IS2 y SECdR_WI;R), como el sistema CRISPR/Cas9 gl₉, se entregaron a las células en forma de plásmido, mediante co-nucleofección. La expresión de eGFP y dsRed se determinó a las 48h y a día 30. b) Gráfico mostrando la eficacia relativa de RH con cada una de las combinaciones DNA donador / sistema CRISPR/Cas9 empleadas. c) Gráficas mostrando la frecuencia relativa de inserciones ilegítimas del donador (sin utilizar RH) con cada una de las combinaciones DNA donador / nucleasa. Los datos representan la media \pm el error estándar de al menos 3 experimentos independientes. La comparación entre los distintos grupos se llevó a cabo mediante el método no paramétrico de Dunn (*p < 0.05).

1.14.1.2 Generación de la línea reportera K562 SEWAS84

Los niveles de especificidad detectados en el modelo anterior eran muy elevados y muy superiores a los datos obtenidos en el locus endógeno²¹⁵. Esto podría ser debido al elevado número de sitios diana (13) encontrados en ese modelo que incrementaban las posibilidades de encontrar su diana en el genoma. Se procedió por tanto a desarrollar otro modelo complementario que tuviera una sola copia del gen diana reportero y que tuviera un método de lectura alternativo. Para ello se diseñó el plásmido lentiviral SEWAS84 que contenía una diana más pequeña del locus *WAS* justo en 5' del gen eGFP intacto, dentro del cassette SE. Con este plásmido se produjeron partículas y se procedió a generar la línea K562 SEWAS84 tal y como se indica en la Figura 36, de manera que nos asegurásemos de tener una población heterogénea pero con una sola integración del vector. El clonaje de esta población dio lugar al clon 1 K562 SEWAS84, con el que procedimos a realizar los siguientes experimentos.

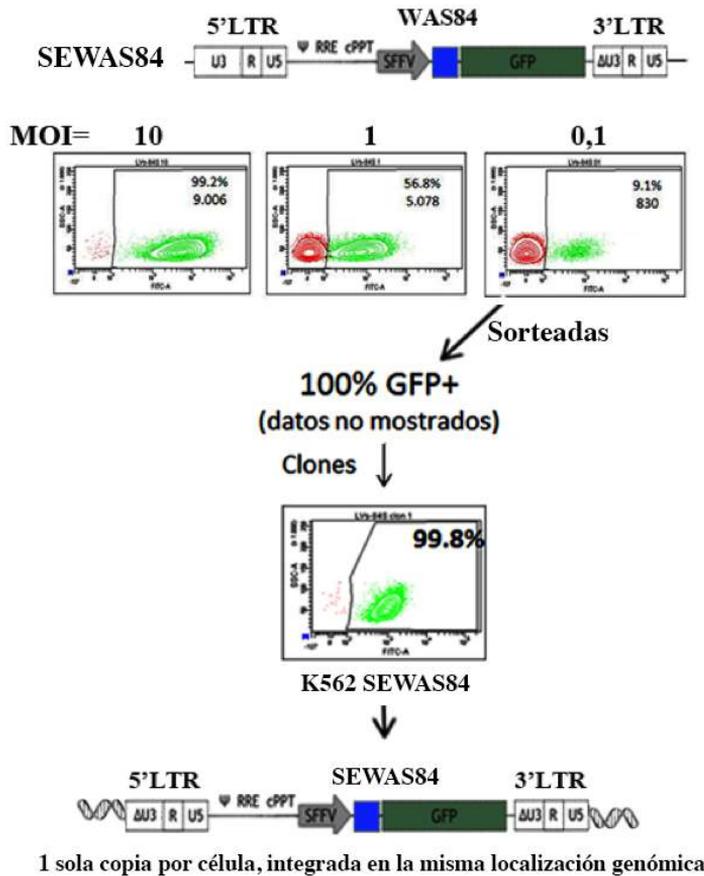


Figura 36. Generación de la línea reportera K562 SEWAS84. a) Esquema representativo del plásmido lentiviral SEWAS84. b) Obtención de la línea K562 SEWAS84 mediante transducción de células K562 con diferentes MOI del vector lentiviral SEWAS84, para conseguir una población con una sola integración del vector en su genoma (plot de la derecha). Se sorteó la población eGFP positiva procedente de la transducción con el MOI más bajo. Esta población se sorteó para la obtención de clones, generándose la línea celular K562SEWAS84 a partir del clon 1.

1.14.1.2.1 Diseño del donador SdR*GFP

Una vez generada la línea SEWAS84, necesitábamos un donador que, de forma sencilla, nos permitiese determinar la eficiencia de RH de nuestros sistemas de EG, a la vez que nos proporcionase un método de cuantificar las inserciones en localizaciones del genoma distintas a la diana. Para ello, diseñamos el plásmido donador SdR*GFP, en el que el promotor SFFV actuaría como brazo de homología en 5', y la secuencia mutada de eGFP lo haría en 3', permitiendo así el silenciamiento de eGFP en aquellas células SEWAS84 en que se produjese RH. Entre ambas secuencias se incluyó el cDNA de la proteína reportera dsRed, que bajo el promotor SFFV proporcionaría la expresión de esta proteína. De esta forma, la expresión de ambos reporteros en células previamente editadas nos indicaría que la integración del donador habría sucedido fuera de la diana, mientras que la expresión de dsRed y la pérdida de eGFP indicarían la integración del donador en el sitio diana (Figura 37).

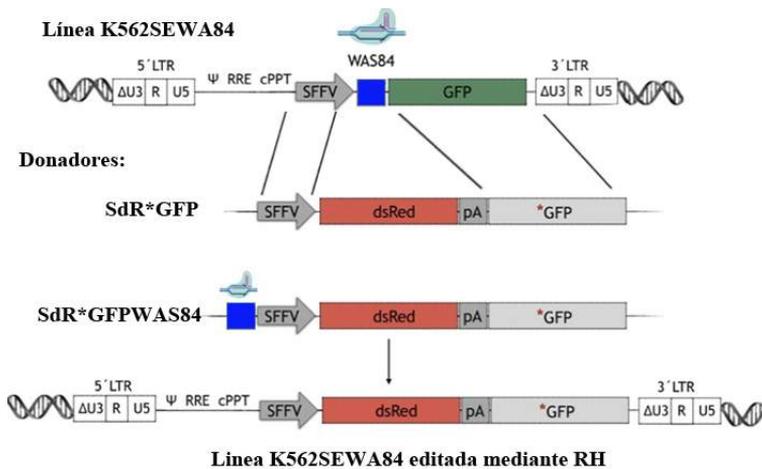


Figura 37. Esquema del modelo K562 SEWAS84 y de los donadores utilizados para analizar eficacia y especificidad. Arriba se ilustra cómo quedaría integrado el cassette diana en el genoma de las células K562 SEWAS84. En medio se describen los dos donadores utilizados sin (SdR*GFP) ó con (dR*GFPWAS84) diana para el gI1 (intrón 1 de *WAS*). Abajo se ilustra como quedaría integrado el donador en caso de RH.

Previamente, en la línea celular K562- δ GF-I1-P habíamos observado que la inclusión en el donador de la diana para el sistema CRISPR/Cas9 aumentaba significativamente la eficacia de RH, sin afectar negativamente a la especificidad. Por ello, sobre la construcción SdR*GFP decidimos, adicionalmente, añadir el fragmento de *WAS* de 84pb, dentro del cual se encuentra nuestra diana para el sistema CRISPR/Cas9 gI1, aguas arriba del promotor SFFV, generando así un segundo donador, denominado SdR*GFPWAS84. La incorporación de ambos donadores en el sitio diana, mediada por RH, supondría en la línea celular SEWAS84 la pérdida de expresión de eGFP, y la ganancia de expresión de dsRed. Si se produce la integración del donador en sitios ilegítimos, pero no en la diana, las células SEWAS84 continuarían expresando eGFP, y pasarían a expresar también dsRed (Figura 38).

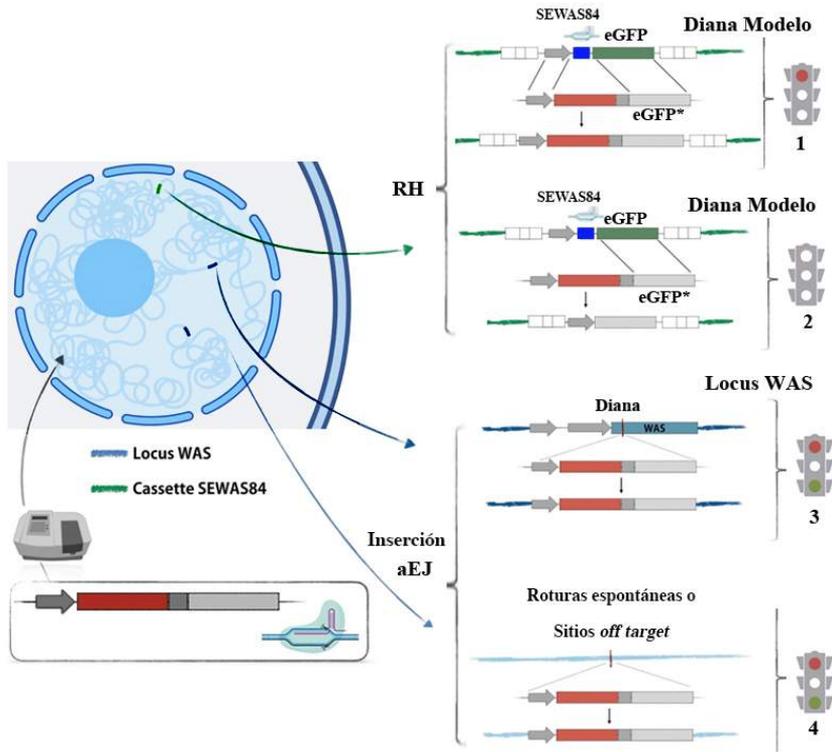


Figura 38. Ilustración esquematizando, de forma simplificada, el tipo de inserciones del donador que se podrían producir en el genoma de la línea celular SEWAS84. En la parte superior izquierda aparece representado el núcleo de células SEWAS84, conteniendo una única copia del cassette SEWAS84 (marcada en verde), y dos copias del locus *WAS* (marcadas en azul). A la derecha se esquematizan los cuatro tipos de integraciones del donador que podrían producirse. Las dos superiores se producirían por RH y las dos inferiores por integración mediante microhomología (aEJ). 1- integración del donador mediante RH en la diana (cassette SEWAS84), que supondría la pérdida de expresión de eGFP y ganancia de dsRED, 2- integración del donador mediante RH en la diana (cassette SEWAS84) pero utilizando solo el brazo de eGFP mutado (eGFP*), que supondría la pérdida de expresión de eGFP, 3- integración del donador en el locus *WAS* endógeno, donde se localiza la diana original del sistema CRISPR/Cas9 g11, y 4- integración del donador en otros sitios ilegítimos del genoma en los que el sistema CRISPR/Cas9 g11 esté produciendo cortes de forma inespecífica (sitios *off target*), o en los que se hayan producido roturas espontáneas del DNA en presencia del donador.

Esta segunda construcción, denominada SdR*GFPWAS84, se compararía a continuación, en términos de eficiencia y seguridad, con SdR*GFP.

1.14.1.3 Comparativa de la eficacia y especificidad de la EG de diferentes donadores

En una primera prueba de concepto, quisimos comparar la eficacia y especificidad de interrupción de la expresión de eGFP en la línea SEWAS84 al emplear como donadores los plásmidos SdR*GFP y SdR*GFPWAS84 (conteniendo este último la diana para el sistema CRISPR/Cas9), y Lgl₁Cas9 en forma de plásmido como método de entrega del sistema CRISPR/Cas9.

Como control negativo, células SEWAS84 se nucleofectaron con cada uno de los plásmidos donadores, en ausencia del sistema CRISPR/Cas9 gl₁. La expresión de eGFP y dsRed comenzó a medirse, mediante citometría de flujo, 48 horas después de proporcionar donador y nucleasa, por nucleofección, a las células SEWAS84, y se midió de forma periódica, hasta su estabilización (Figura 39).

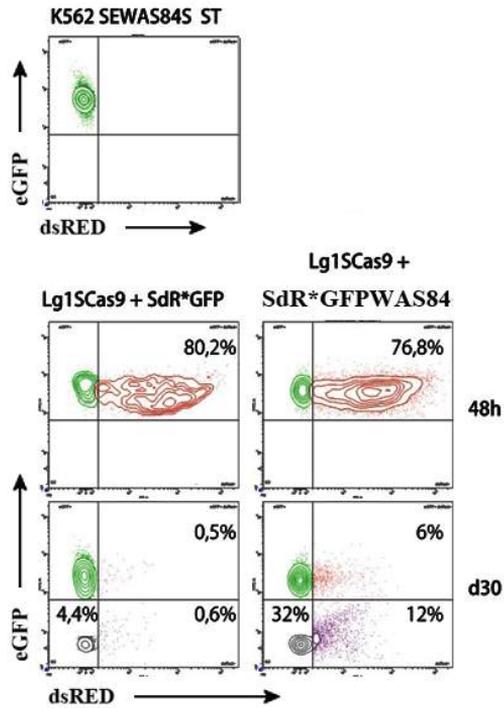


Figura 39. Silenciamiento de eGFP en línea celular reportera K562 SEWAS84, empleando distintos donadores. Plots de citometría representativos mostrando seguimiento en el tiempo de la expresión, tanto de eGFP (FITC-A), como de dsRed (PE-A). El plot superior muestra las células del modelo sin transfectar. Los plots centrales muestran los niveles de eGFP y dsRed en las células K562 SEWAS84 nucleofectadas con Lg1SCas9 plásmido en combinación con el donador SdR*GFP (izquierda), o con su versión alternativa conteniendo la diana para Lg1SCas9, el donador SdR*GFPWAS84 (derecha), 48h después de la nucleofección. Los plots inferiores muestran los niveles de expresión de eGFP y dsRed en estas células determinados a día 30.

La eficiencia relativa de RH se determinó a tiempo final, relativizando el porcentaje de células eGFP-/dsRed+ en este punto, respecto a la eficacia de entrada del donador en las células diana, definida como el porcentaje de células dsRed positivas 48h después de la nucleofección. Al valor resultante se sustrajo el obtenido en células a las que se había suministrado el correspondiente plásmido donador en ausencia del sistema CRISPR/Cas9 gI1.

Como se aprecia en la Figura 40, tanto el porcentaje de RH obtenido en cada uno de los experimentos independientes, como la eficacia relativa de RH, considerando el conjunto de los datos, fue bastante superior al emplear como donador el plásmido SdR*eGFPWAS84, conteniendo la diana para Lg1Cas9.

Por otro lado, para determinar la especificidad de integración de cada uno de los donadores en el modelo K562 SEWAS84, fue necesario tener en cuenta que la eficiencia de RH esta estimada a la baja, ya que un porcentaje elevado de las células eGFP-dsRED- pueden ser debidas a RH con el donador, pero solo en la región de eGFP mutada. También hemos observado un cierto grado de silenciamiento de eGFP en estas células cuando se les suministra el sistema CRISPR/Cas9 sin donador, lo que podría ser debido a la generación de indels por NHEJ que afecten al primer ATG de eGFP (dado que la diana del sistema CRISPR/Cas9 se sitúa a tan solo 36pb de este primer ATG), o al estrés al que se ven sometidas estas células. Por ello, teniendo en cuenta que mediante citometría no podemos establecer qué porcentaje de la población doble negativa corresponde a eventos de RH, decidimos analizar la relación entre integraciones ilegítimas y especificidad de RH considerando dos supuestos: 1- las células dobles negativas no proceden de eventos de RH, 2- todas las células dobles negativas proceden de RH. De esta forma, obtendríamos el rango dentro del cual se encontraría nuestro sistema CRISPR/Cas9. Así, obtuvimos que la relación entre el grado de integraciones ilegítimas y de eficiencia de RH se situó entre un 1,8% y un 48% en el

caso del donador SdR*GFP, y entre un 13% y un 42% en el caso del donador SdR*GFPWAS84, siendo este último, por tanto, el que estaría produciendo una mayor proporción de integraciones en sitios ilegítimos con relación a inserciones dentro de la diana. No obstante, este método de análisis no nos permite establecer un valor exacto.

En el modelo K562- δ GF-I1-P, el cual contiene 13 copias del cassette SE- δ GF-I1-P, los niveles de integración del donador en sitios ilegítimos se situaron siempre muy por debajo del 1%. Esto podría deberse a que existe una competencia entre los sitios *off target* y la diana por la unión de Cas9, donde predomina la unión con su diana. Al contar la línea celular K562- δ GF-I1-P con 13 copias de la diana en su genoma, la RH estaría viéndose favorecida, y estaríamos sobrestimando la especificidad de nuestras herramientas. Por otro lado, el modelo SEWAS84, al contener una única copia del cassette SEWAS84, nos estaría proporcionando datos de especificidad y eficiencia de RH más representativos de lo que ocurriría en el locus *WAS* endógeno.

Cabe destacar que la relación entre inserciones fuera de la diana y frecuencia de RH fue similar con ambos donadores, de forma que, aunque con el donador WAS84SdR*GFP se producen un mayor número de integraciones ilegítimas, este incremento es proporcional al incremento en el número de integraciones en la diana. Es decir, el donador SdR*GFPWAS84 incrementa la eficacia de RH sin producir un efecto negativo sobre la especificidad.

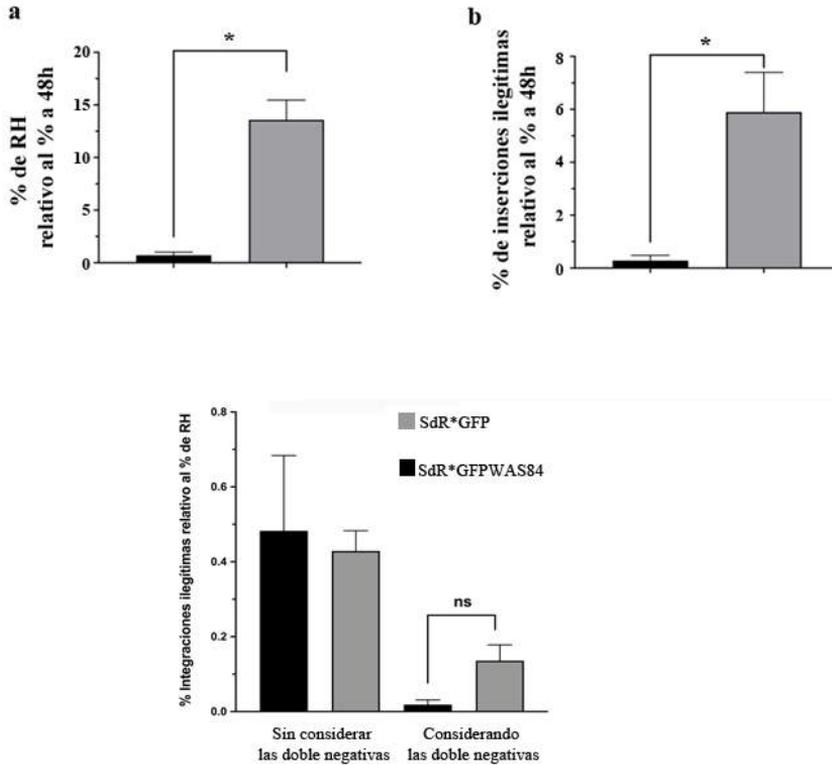


Figura 40. Silenciamiento de eGFP en línea celular reporteraK562 SEWAS84, empleando diferentes donadores y el sistema CRISPR/Cas9 gl_1 , entregados como plásmidos. a) Gráfico mostrando la eficacia de RH con cada uno de los donadores, relativa a la eficiencia de nucleofección. La eficacia relativa de RH para cada una de las condiciones se calculó relativizando el porcentaje de células eGFP negativas y dsRed positivas a tiempo final con respecto al porcentaje de células dsRed positivas 48h post-nucleofección, y extrayendo el valor obtenido para el control negativo (células nucleofectadas con el correspondiente plásmido donador, en ausencia del sistema CRISPR/Cas9 gl_1). b) Gráficas mostrando la frecuencia de inserciones ilegítimas del donador (fuera de la diana o inserciones del DNA donador completo en la diana) en cada uno de los experimentos independientes (izquierda), y la frecuencia relativa de este tipo de inserciones en el conjunto de los datos (derecha). La frecuencia relativa de inserciones ilegítimas se determinó en base a eficiencia inicial de entrada del donador en las células diana, corrigiendo respecto al valor obtenido en células nucleofectadas con el

Resultados

correspondiente plásmido donador en ausencia de nucleasa. Las gráficas muestran datos de tres experimentos independientes. c) Relación entre integraciones ilegítimas y eficiencia de RH, relativas a eficiencia de nucleofección. Este análisis se ha realizado considerando dos posibilidades: 1-población de células dobles negativas no procede de eventos de RH (gráficas izquierda), 2-población de células dobles negativas procede de eventos de RH (gráficas derecha). Los datos representan la media \pm el error estándar de al menos 3 experimentos independientes. La comparación entre los distintos grupos se llevó a cabo mediante la prueba U de Mann-Whitney (* $p < 0.05$).

1.15 OBJETIVO 3: EFICIENCIA Y ESPECIFICIDAD DE SISTEMAS CRISPR/Cas9 PARA *WAS* EN CÉLULAS T Y HSCs

Dado que la nucleofección de RNP era el método más eficaz para generar indels en células T y HSCs, procedimos a diseñar estrategias para la EG del locus *WAS* mediante RH.

1.15.1.1 Eficacia y seguridad de corte en el locus *WAS*

En primer lugar, procedimos a analizar la eficacia de corte de diferentes gRNAs dirigidos frente al primer ATG del gen *WAS*, con la finalidad de detectar un cambio en la expresión de esta proteína en células editadas. Se diseñaron dos gRNAs dirigidos frente al primer exón de *WAS* (gE₁ y gE₂), y evaluamos la disrupción del gen *WAS* en células T y HSCs, usando estos sistemas (RNP_gE₁ y RNP_gE₂) (Figura 41).

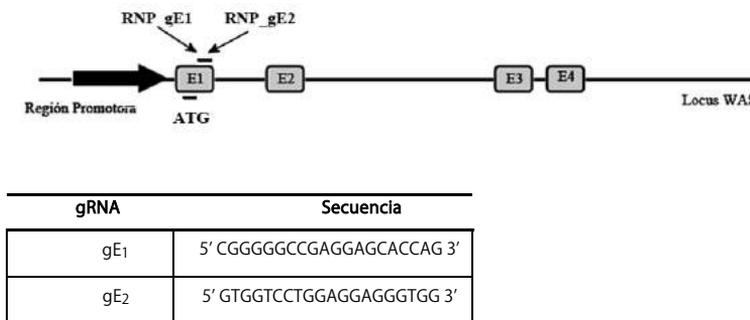


Figura 41. Esquema de las dianas de los gRNAs frente al exón 1 del locus *WAS* localizados justo 3' del ATG, y tabla mostrando cada uno de los gRNAs.

Los datos obtenidos confirmaban la eficacia observada previamente con nucleofección de RNP con gRNAs frente al primer intrón. Efectivamente, tanto en células T (Figura 42), como en HSCs (Figura 43), ambos gRNAs eran capaces de editar el locus *WAS*, aunque el gE₁ fue, en términos de eficiencia, muy superior al gE₂. Esto se analizó por T7 (Figura 42, a, izquierda), mediante análisis ICE (Figura 42, a, derecha; y Figura 43, a), y mediante la disminución de los niveles de proteína WASP (Figura 42, b, y Figura 43, b).

En células T, la eficacia de corte del sistema RNP gE₁ estimada mediante el ensayo de la T7 endonucleasa difería bastante con respecto a la eficacia estimada mediante el algoritmo ICE, poniendo de manifiesto, de nuevo, que la segunda aproximación es más robusta.

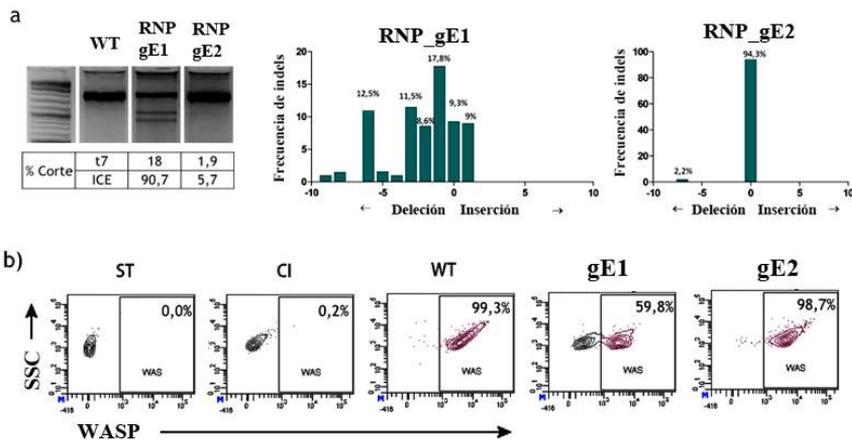


Figura 42. Eficacia en células T de RNP dirigidas frente a la región codificante de *WAS*. a) Eficacia de corte de sistemas CRISPR/Cas9 dirigidos frente al primer exón del locus *WAS*, proporcionados a células T como RNP. Esta eficacia de corte se determinó por el ensayo de la T7, así como mediante el algoritmo ICE. Los gráficos de la derecha muestran el patrón del tipo de indels generados en el locus

WAS con RNP gE₁, o RNP gE₂. b) Análisis por citometría de flujo de la expresión de *WASP* en células T (ST: sin teñir, CI: control de isotipo, WT: células T no editadas, gE₁: células T editadas con RNP gE₁, gE₂: células T editadas con RNP gE₂).

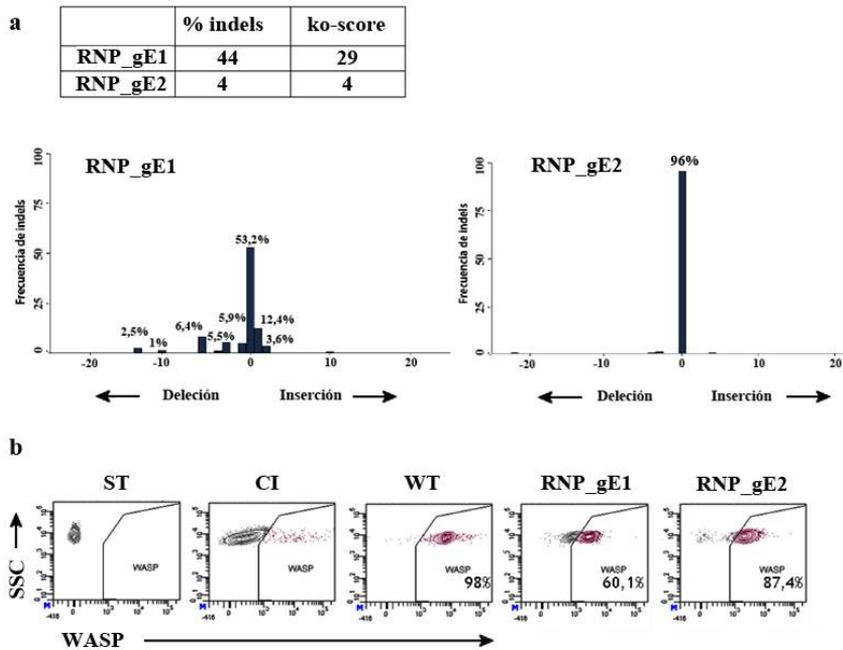


Figura 43. Eficiencia en HSCs de RNP dirigidas frente a la región codificante de *WAS*. a) Eficacia de corte de sistemas CRISPR/Cas9 dirigidos frente al primer exón del locus *WAS*, proporcionados a HSCs como RNP. Esta eficacia de corte se determinó, en HSCs, mediante el algoritmo ICE (ver material y métodos). La tabla muestra el porcentaje de indels generados en el locus *WAS* con RNP gE₁, o RNP gE₂ (primera columna), así como el porcentaje de estos indels que provocaría un cambio en el marco de lectura. Las gráficas inferiores muestran el patrón del tipo de indels generados en el locus *WAS* con RNP gE₁, o RNP gE₂. b) Análisis por citometría de flujo de la expresión de *WASP* en HSCs (ST: sin teñir; CI: control de isotipo; WT: HSCs no editadas; RNP gE₁: HSCs editadas usando el gE₁; RNP gE₂: HSCs editadas usando el gE₂).

Tal y como se hizo con el g₉, quisimos estudiar si el guía que había mostrado una mayor eficacia para la interrupción del primer exón de *WAS* en células T y HSCs generaba, en estas células, indels en los principales sitios *off target* propuestos por herramientas de predicción *in silico*. Con esta finalidad, se utilizó la herramienta online Cas-OFFinder para la predicción de potenciales sitios *off target* para el gE₁, y se seleccionaron los 6 primeros (EOT1-EOT6). A partir de genómico de células T y HSCs editadas con RNP gE₁, se secuenciaron estos sitios y se determinó, en cada uno de ellos, el porcentaje de indels mediante ICE.

En células T, RNP gE₁ produjo, en dos de los sitios *off target* propuestos por la herramienta Cas-OFFinder, más de un 10% de indels (Figura 44). En el caso de HSCs, coincidiendo con células T, RNP gE₁ generó en E_OT3 indels, con una frecuencia del 6,4% (Figura 44). Estos datos, aún considerando el sesgo del análisis, sugieren que el gE₁ no es apropiado, por ser inespecífico. En este sentido, se habría de realizar un estudio más sistematizado para encontrar aquellos gRNAs que permitan mejores eficacias y menores *off targets* en células T y en HSCs, antes de proseguir con el diseño de una estrategia de EG para el tratamiento de WAS.

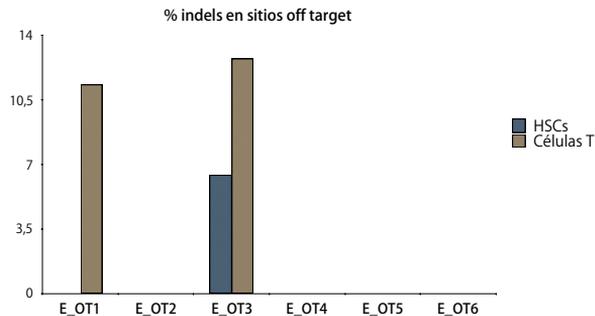


Figura 44. Efectos *off target* del sistema RNP gE₁ en células T y HSCs. Porcentaje (%) de indels generados en potenciales sitios *off target* para el gE₁ (E_OT1 – E_OT6), determinados con el algoritmo ICE. Este análisis se realizó a partir de DNA genómico de células T y HSCs a las que se había entregado, mediante nucleofección, el sistema RNP gE₁ dirigido frente al primer exón de *WAS*.

1.15.2 Eficacia y seguridad de RH en el locus *WAS*

A pesar de los niveles de *off targets* detectados, en la última fase de esta tesis se planteó como último objetivo estudiar la eficacia y seguridad de la EG mediante RH del locus *WAS* en HSCs.

El primer paso para la consecución de este objetivo fue el diseño de un donador con brazos de homología frente al locus *WAS*. El brazo de homología en 5' se diseñó conteniendo las 760pb de la secuencia de *WAS* inmediatamente anteriores a la diana del sistema CRISPR/Cas9 gI₉, mientras que el brazo de homología en 3' contendría las 760pb posteriores, excluyendo del diseño la secuencia diana. Entre ambos brazos se localiza el cassette de expresión SE, incluyendo la señal de poladenilación del gen bGH. A raíz de los resultados obtenidos en los modelos celulares para ensayos de RH, K562-

δ GF-I₁-P y SEWAS84, decidimos incluir en este donador, aguas arriba del brazo de homología en 5', y en orientación inversa, la diana para el sistema CRISPR/Cas9 gl₉, para favorecer la RH en HSCs (Figura 45).

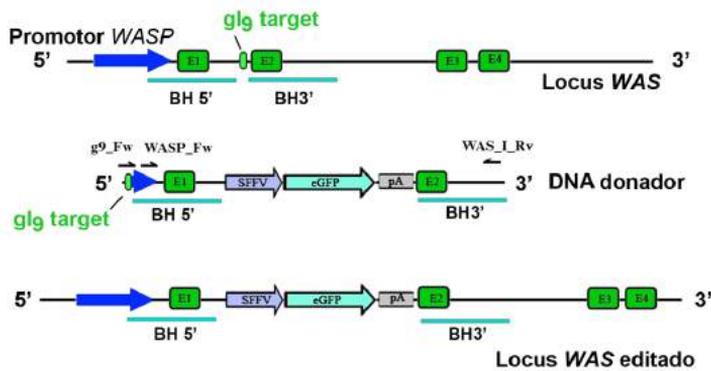


Figura 45. Edición del locus *WAS* mediada por RH. En el diagrama superior de la figura se representan los primeros 4 exones del locus *WAS* (E1-E4), la secuencia diana para el sistema CRISPR gl₉, en el primer intrón de *WAS*, y las secuencias usadas para promover la RH (BH 5' y BH 3'). El diagrama central muestra el DNA donador usado para editar el locus *WAS*. Entre ambos brazos de homología se ha insertado el gen reportero *eGFP* bajo el promotor SFFV. Precediendo al brazo de homología 5' se localiza la diana para el sistema CRISPR gl₉, en orientación invertida. Los primers utilizados para obtener los donadores mediante PCR con (g9_Fw/ WAS_I_Rv) o sin (WASP_Fw / WAS_I_Rv) sitio diana se indican con flechas. El diagrama inferior muestra cómo quedaría el locus *WAS* tras la incorporación del DNA donador mediada por RH. La introducción de este donador en el locus *WAS* suprime la expresión de la proteína WASP, debido a la presencia de la señal de poladenilación del gen bGH.

En una primera aproximación, decidimos entregar el donador pUC57 WAS-SE-g9 a HSCs en forma de plásmido, junto al sistema CRISPR/Cas9 gl₉, entregado como RNP (Fig 35a, plots de la izquierda). Además, decidimos testar la entrega del donador en forma de producto de PCR, tras comprobar que algunos autores proponen este sistema de entrega. Para la entrega del DNA donador como producto de PCR se

diseñaron dos parejas de primers. La primera de ellas (g9_Fw / WAS_I_Rv) permitiría la amplificación del DNA donador incluyendo la diana para el sistema CRISPR/Cas9 g₉ (WAS-SE-g9), mientras que la segunda (WASp_Fw / WAS_I_Rv) produciría un fragmento similar, WAS-SE, en el que se hubiese perdido la diana.

Los distintos donadores (pUC57 WAS-SE-g9, y los productos de PCR WAS-SE-g9 y WAS-SE) se entregaron a HSCs de varón en d6 post-aislamiento, junto al sistema CRISPR/Cas9 g₉ como RNP, mediante nucleofección. Como control, también se nucleofectaron HSCs con cada uno de los donadores, en ausencia de RNP. La entrega del donador en forma de producto de PCR (con o sin diana) se llevó a cabo empleando cantidades equimolares a las cantidades usadas de plásmido donador (2µg).

Nuestros resultados mostraron que la entrega del donador como producto de PCR no solo fue más eficiente que como plásmido, sino que fue la única de las dos estrategias que en la que se observó RH (establecimiento de una población eGFP positiva estable) (Figura 46). Por otro lado, a diferencia de lo observado en nuestros modelos celulares, aunque el donador WAS-SE-g9 fue ligeramente más eficaz que WAS-SE, las diferencias no fueron significativas. Por lo que, en este caso, no podemos concluir que la inclusión de la diana favorezca la RH en HSCs. Sin embargo, este experimento necesita repetirse más veces y con diferentes concentraciones de donador para poder llegar a una conclusión en HSCs.

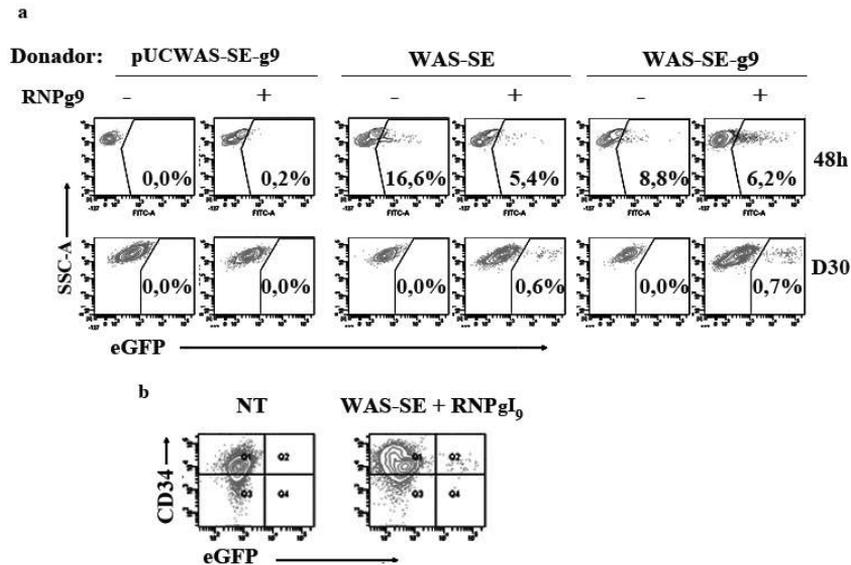


Figura 46. Evaluación de la eficiencia de RH en células CD34⁺ al entregar el sistema CRISPR/Cas9 gI9 como RNP, y el donador para *WAS* como plásmido, o como producto de PCR. a) Plots de citometría mostrando eficiencia de entrada del donador en HSCs a 48h (plots superiores), y la eficiencia de RH (plots inferiores) estimada como el porcentaje de células eGFP positivas 30 días después (d30). Se compara la eficacia de donadores en forma de plásmido (pUCWAS-SE) y de productos de PCR (WAS-SE-g9 y WAS-SE) b) Plots representativos de la expresión de CD34 en HSCs editadas y no editadas, determinada mediante tinción con anti-CD34 23d después de la nucleofección.

La expresión de eGFP en HSCs, una vez diluida la expresión episomal del DNA donador, debería ser consecuencia de la integración del DNA donador en el locus *WAS*. Una forma de comprobar esto sería estudiar la expresión de WASP en células que hubiesen incorporado el DNA donador en su genoma. Decidimos analizar la expresión de WASP en células que habían incorporado el DNA donador, para comprobar si de esta forma lográbamos distinguir integraciones del donador en el locus *WAS* de integraciones en

sitios inespecíficos. Este análisis se llevó a cabo sobre HSCs a las que se había proporcionado, junto a las RNP, el DNA donador como producto de PCR incluyendo la diana para el sistema CRISPR/Cas9 gI9. Los niveles de expresión de WASP se determinaron en la subpoblación negativa para la expresión de eGFP (P3, en la Figura 47), y que por tanto no había incorporado el donador en su genoma, así como en la subpoblación que expresaba eGFP, en la que el donador se habría integrado (P2, en la Figura 47). Debido a limitaciones de la técnica, no fuimos capaces de distinguir dos poblaciones claras para la expresión de WASP en las células eGFP positivas. Sin embargo, si pudimos apreciar una disminución en los niveles de expresión de WASP en las células eGFP positivas, tanto en porcentaje (3.75% en Q1 frente a 0,02% en Q2) como en MFI (4724 de la población eGFP negativa, frente 3805 de la población eGFP positiva). Esto indicaba que el donador se habría integrado, al menos en parte, en el locus *WAS*. No obstante, aún se requiere una optimización del protocolo de tinción de WASP para poder distinguir las integraciones del donador dentro de la diana de aquellas que hubiesen ocurrido en localizaciones diferentes a esta.

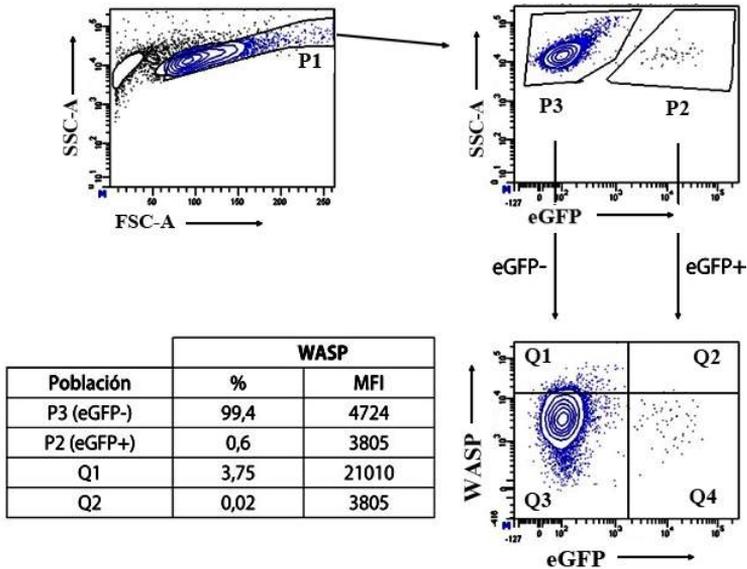


Figura 47. Análisis de la expresión de WASP en HSCs que han incorporado el DNA donador en su genoma. Los niveles de expresión de WASP se analizaron en HSCs procedentes de varón, a las que se había proporcionado el sistema CRISPR/Cas9 gI9 como RNP, y el DNA donador para el gen *WAS* como producto de PCR incluyendo la diana para el gI9. Este análisis se llevó a cabo mediante tinción intracelular de WASP y citometría de flujo. Los niveles de expresión de WASP (MFI en APC) se analizaron en las subpoblaciones negativa (P3, en plot superior, derecha) y positiva para la expresión de eGFP (P2, en plot superior, derecha). Los niveles de expresión de WASP de cada una de las subpoblaciones aparecen indicados en la tabla que se muestra en la figura.

DISCUSIÓN

El gran impulso que está sufriendo las técnicas de edición genómica, particularmente desde la irrupción del sistema CRISPR/Cas9 en este campo, acerca a la realidad la posibilidad de abrir nuevas vías para el tratamiento de determinadas enfermedades que actualmente solo cuentan con tratamientos dirigidos a paliar su sintomatología. La aplicación de estas técnicas, que permiten modificar secuencias del genoma de forma selectiva y precisa, resultan de especial interés para el tratamiento de enfermedades monogénicas recesivas, donde la inserción del gen terapéutico en el genoma de la célula diana, o la corrección mediante edición génica de la mutación causante de la enfermedad, favorecería la restauración del fenotipo sano en pacientes, de una forma permanente. Dentro de este grupo de enfermedades para las que la EG podría representar una alternativa terapéutica, se encuentra el Síndrome de Wiskott Aldrich, inmunodeficiencia primaria causada por mutaciones en el gen *WAS*, donde el único tratamiento curativo es el trasplante alogénico de médula ósea. La restauración de la expresión de la proteína WASP en células hematopoyéticas de pacientes WAS favorecería una desaparición, al menos parcial, de las manifestaciones clínicas, en casos en que no exista la posibilidad de un trasplante alogénico, o exista riesgo.

Las principales células diana frente a las que iría dirigida una posible estrategia de EG para WAS serían las HSCs de paciente, dado que la diferenciación de HSCs en las que se hubiese restaurado previamente la secuencia de *WAS* favorecería que células de todos los tipos celulares de la sangre portasen esta corrección y, por tanto, expresasen la proteína WASP. Con la modificación genética de HSCs se puede lograr, no sólo la reconstitución inmunológica de pacientes inmunodeficientes (SCID-X1, WAS, ADA), sino que también permite la expresión de proteínas necesarias para el rescate de otros fenotipos no inmunes (MLD, ALD), ya que cualquier modificación en el genoma de HSCs se transmitirá a los distintos linajes hematopoyéticos. Sin embargo, las HSCs son probablemente unas de las células más difíciles de modificar. Además, su aislamiento y expansión alteran su fenotipo y reducen su capacidad de diferenciación.

Otra posible diana es la célula T, donde la expresión de WASP es esencial para su normal desarrollo, proliferación y funcionalidad. Además, su modificación resulta más sencilla en comparación con las HSCs, presentan un menor riesgo de transformación, pueden aislarse en gran número de forma eficiente, y pueden permanecer en los pacientes durante años. Estas células ya han mostrado su potencial como diana terapéutica, en inmunoterapia de leucemias y linfomas, principalmente para la generación de células T expresando receptores de antígenos quiméricos (células CAR T), así como para el tratamiento de infección por VIH, donde ya se ha conseguido la disrupción de forma eficiente el receptor CCR5 mediante distintas herramientas de edición génica (ZFNs, TALEN, y CRISPR), confiriendo a estas células T resistencia frente a la infección por el virus.

Dos aspectos fundamentales que debe garantizar una estrategia de EG que persiga la modificación de células primarias, como HSCs, o células T, son la entrada de las herramientas de EG (nucleasas y/o donadores) al núcleo de estas células de forma muy eficiente, y la bioseguridad de las células diana, que no deben ver afectada su funcionalidad. Por ello, antes de que una estrategia de EG llegue a ensayos clínicos, es vital ahondar en investigar su eficacia terapéutica, así como su bioseguridad, tanto en modelos celulares, como en las células diana.

Teniendo en cuenta todo lo expuesto, en esta tesis se planteó desarrollar herramientas para la EG eficaz y segura en células T y HSCs. Para ello, en primer lugar se han comparado diferentes sistemas de entrega, tanto de las NE como de los DNA donadores. Para poder estudiar en detalle la seguridad de los sistemas, se han generado dos modelos celulares donde poder analizar de una manera simple la especificidad de EG. Finalmente, se han analizado los mejores sistemas disponibles para la edición del gen *WAS* en células HSCs y células T.

Evaluación de la eficiencia y especificidad de corte en el locus *WAS* utilizando diferentes sistemas de entrega en células K562

Si bien HSCs y células T procedentes de paciente serían las dianas frente a las que se dirigiría una posible estrategia de EG para *WAS*, la disponibilidad de este tipo de células para ensayos preliminares es muy limitada, dada la baja frecuencia de la enfermedad, así como la complejidad de obtención de estas células. La disponibilidad tanto de HSCs como de células T procedentes de donante sano son también limitada, y la investigación con ellas ha de ser limitada a los estadios finales, tras la optimización de las herramientas en células más fáciles (y baratas) de obtener y manipular. Por lo tanto, decidimos optimizar la entrega de nuestros sistemas CRISPR/Cas9 específicos para *WAS*, previamente, en la línea celular K562, de origen hematopoyético. En estas células se evaluó la eficacia y especificidad de dos sistemas CRISPR dirigidos frente al primer intrón del gen *WAS* (sistema CRISPR/Cas9 g1 y sistema CRISPR/Cas9 g1₉), al ser vehiculizados hasta las células diana siguiendo tres estrategias distintas: entrega en forma de plásmido, de IDLV, o de RNP, todos ellos sistemas transitorios. Era esperable que los niveles de Cas9 y/o gRNA por célula fuera diferente con cada uno de los sistemas de transferencia y esto pudiese dar lugar a diferencias en efectos *off-target*. En este sentido, nuestra hipótesis era que la utilización de IDLVs podría proporcionar una expresión óptima y transitoria de Cas9 y gRNA en las células diana en función del MOI utilizado y/o del diseño del vector. Sin embargo, en K562, los tres métodos de entrega mostraron eficacias y niveles de *off-targets* similares. Aunque tal y como ha sido descrito por otros autores, la nucleofección con RNP parecía el método más robusto. Hasta el momento, este método de entrega ha demostrado ser el más eficaz para la entrega de Cas9 en distintos tipos celulares, proporcionando altos niveles de proteína con una expresión muy transitoria, al ser entregada ya en forma de ribonucleoproteína preformada. Por otro lado, si observamos un patrón de generación de indels diferente en función del método de entrega del sistema CRISPR/Cas9,

aunque estos datos tendríamos que confirmarlos incrementando el número de experimentos y analizando *off-targets* de manera no sesgada.

Evaluación de la eficiencia y especificidad de corte en el locus *WAS* utilizando diferentes sistemas de entrega en células T y HSCs

Dado que los tres sistemas ofrecieron niveles similares de eficacia y especificidad en K562, procedimos a analizar los tres en HSCs y células T. Sin embargo, al contrario que en K562, el único método de entrega que nos permitió ver edición del locus *WAS* en ambos tipos celulares fue la entrega en forma de RNP. Han sido varios los grupos que, como nosotros, han reportado la falta de solidez en los protocolos de modificación de HSCs²⁵⁰, si bien es verdad que también ha habido grupos que han publicado resultados prometedores utilizando IDLVs. En cualquier caso, la nucleofección con RNP está siendo el método de elección para la EG tanto de células T, como de HSCs^{251–253}. A pesar de estos datos, teníamos varios indicios que nos hacían pensar que, si incrementáramos la eficacia de los IDLVs, se podría conseguir desarrollar un sistema de entrega de CRISPR/Cas9 ideal, dado que podríamos controlar los niveles de expresión de Cas9 y gRNA. Además, la mejora de los IDLVs podría tener diversas aplicaciones, al margen de la EG.

Desarrollo de nuevos sistemas de entrega más eficaces basados en IDLVs.

Los IDLVs representan una herramienta de entrega atractiva porque mantienen las principales propiedades de los vectores lentivirales integrativos, pero sin integrarse en el genoma de la célula hospedadora. Son, por tanto, más seguros cuando se requiere una expresión estable en células que no se dividen (neuronas, hepatocitos), pero sobre todo, y desde el punto de vista de esta tesis, también pueden lograr una expresión transitoria en células en división activa. Sin embargo, los niveles de expresión y los títulos de los IDLV son generalmente más bajos que los de los LV, que ha limitado las aplicaciones de esta tecnología.

Pese a que en células T y HSCs no conseguimos editar el gen *WAS* utilizando IDLVs, teníamos la hipótesis de que si lográbamos mejorar los IDLVs como plataforma de entrega, conseguiríamos los niveles de Cas9 y gRNA necesarios y suficientes (umbral) para editar estas células de forma eficaz y segura. Pensábamos que esto podría reducir la generación de cortes en sitios *off target*. Por ello, nos propusimos trabajar en el desarrollo de IDLVs mejorados, que podrían emplearse, además, para otras aplicaciones en las que se requiera una expresión transitoria, o para expresión estable en células quiescentes, evitando la integración del transgén. Lamentablemente, aunque conseguimos optimizar el funcionamiento de los IDLVs en varios tipos celulares, en HSCs y en células T no se consiguió dicho objetivo.

Los niveles de expresión relativamente bajos de los IDLV se han relacionado con el silenciamiento epigenético a través de mecanismos de defensa celular que aplican marcas de heterocromatina a las secuencias virales episomales. Esta respuesta celular no está restringida a IDLVs, sino que afecta a otro tipo de vectores, como a vectores derivados del virus del herpes simple y los Adv. El DNA viral episomal está "cromatinizado" y adquiere propiedades similares a los nucleosomas inmediatamente después de la entrada en el núcleo. En particular, se ha informado previamente que los genomas IDLV se someten a heterocromatinización a través de la desacetilación de histonas, un proceso que se puede revertir usando inhibidores de desacetilación de histonas (HDACi) como el butirato de sodio y el ácido valpróico. Para investigación básica y algunas aplicaciones clínicas, HDACi podría utilizarse para mejorar la eficiencia de IDLVs; sin embargo, en la mayoría de las estrategias de TG, el uso de HDACi no es deseable debido a posibles efectos secundarios graves, como el desarrollo de tumores malignos. Previamente, en nuestro laboratorio se había diseñado un elemento aislador (IS2), que combina un elemento SAR sintético (SAR2) y un fragmento de 650pb del aislador cHS4 del locus de la globina de pollo la globina de pollo (HS4-650), que en

vectores lentivirales había mostrado la capacidad de reducir la variabilidad de la expresión y mejorar los niveles de expresión en células de gran interés, como son hESC indiferenciadas y células hematopoyéticas derivadas de hESC. Nos planteamos, por tanto, la hipótesis de que la inclusión del elemento IS2 en los IDLV podría mejorar su comportamiento al evitar el silenciamiento epigenético a través de la actividad de HS4 y al mejorar la eficiencia de la transcripción a través de la actividad de SAR.

Contrariamente a lo esperado, la presencia del elemento IS2 no anuló el silenciamiento epigenético de las HDAC, aunque mejoró la eficiencia transcripcional de los episomas. Aunque la presencia de secuencias SAR en el elemento IS2 también podría afectar la estabilidad de la expresión, tampoco encontramos este tipo de efecto, que está en línea con las observaciones de Kymäläinen et al., que no observaron diferencias en la estabilidad episomal en IDLVs conteniendo una secuencia SAR. Sin embargo, estos datos difieren de otros estudios que muestran que la inserción de un fragmento completo de 1.2 kb o mínimo de 155 pb de los elementos SAR del interferón β en IDLVs proporciona una expresión de transgén episomal sostenida, al menos en una parte de la población. Estos hallazgos contradictorios podrían ser explicados por las diferencias en los elementos SAR utilizados en los diferentes IDLV.

El elemento IS2 se insertó en el 3'LTR del esqueleto lentiviral para duplicarlo durante el proceso de transcripción inversa. Esto mejora el efecto de las secuencias SARs (cuantas más secuencias mejor) y / o aumenta la recombinación homóloga para promover la formación de círculos 1-LTR, que ha descrito que son de 2 a 4 veces más eficaces para la expresión que círculos con 2-LTR. Sin embargo, nuestro análisis mostró que la inclusión, tanto del elemento IS2, como de una secuencia irrelevante, no afectaba el título del vector pero si redujo la cantidad de episomas en las células. Además, la expresión mejorada de los episomas IS2-IDLV tampoco puede explicarse

por el incremento en las formas 1-LTR, ya que este efecto no está presente en los IDLV SE-1.2kb.

Nuestros datos indicaban que el mejor rendimiento de los episomas que contienen IS2 se deben a un efecto relacionado con la transcripción y no a la estabilidad del mRNA y / u otros efectos en el mRNA. Por lo tanto, centramos nuestra atención en tratar de comprender el mecanismo potencial involucrado en el incremento de actividad transcripcional que estimamos era de 6 a 7 veces mayor que en los episomas de SE que en los SE-IS2. Estos efectos positivos de la secuencia IS2 contrarrestan el efecto negativo en la generación de episomas en las células diana. Por lo tanto, el efecto final del elemento IS2 en los IDLV dependerá en gran medida de la célula diana y del equilibrio entre los efectos negativos (menos eficacia de la generación del episoma) frente a los efectos positivos (transcripción mejorada) del elemento IS2 en cada tipo de célula.

Estudios de posicionamiento nuclear de los episomas SE y SE-IS2 mostraban que los SE-IS2-IDLV seguían un patrón más agregado en las regiones con menor densidad de DAPI, indicando que el incremento la actividad transcripcional de los episomas SE-IS2 probablemente se deba a un reposicionamiento nuclear preferencial en regiones transcripcionalmente activas.

Dado que el efecto neto de la inclusión del IS2 en los IDLVs dependía del balance entre efectos negativos (menos episomas) y efectos positivos (mayor actividad transcripcional), observamos una alta variabilidad en el efecto neto en diferentes líneas celulares. Observamos en efecto neto positivo en células K562, 293T, NPC, neuronas diferenciadas y en las iPSC, mientras que el efecto fue nulo o negativo en otros tipos celulares, incluidas células T y HSCs. Se requieren, por tanto, mejoras adicionales, como la reducción del tamaño del elemento IS2 y / o el diseño de nuevas

esqueletos lentivirales para que los IDLV puedan ser un sistema eficaz de transferencia del sistema CRISPR/Cas9 para su aplicación a EG de nuestras células diana.

Líneas reporteras de RH

Otro de los objetivos de esta tesis era la generación de modelos celulares que nos permitiesen analizar eficiencia y seguridad de corrección génica mediante RH, utilizando sistemas CRISPR/Cas9 específicos para el gen *WAS*. Para ello diseñamos, en primer lugar, el modelo celular K562 SE- δ GF-I₁-P, conteniendo integradas en su genoma en torno a 13 copias del cassette de expresión SE- δ GF-I₁-P, que no expresa eGFP y nos permite analizar la eficacia de rescate de eGFP utilizando nucleasas dirigidas frente al locus *WAS*. Además, empleando el donador adecuado, permite hacer estudios comparativos de los niveles de integraciones ilegítimas del donador (producidas en sitios distintos a la diana).

Utilizando este modelo no pudimos detectar diferencias significativas entre los tres métodos de entrega del sistema CRISPR/Cas9 comparados (RNP, IDLVs, plásmidos), ni en la seguridad ni en especificidad. Sorprendentemente, encontramos unos niveles muy altos de especificidad (por encima del 99% de integración del DNA donador en el sitio diana, mediante RH). Estos datos diferían con datos anteriores del grupo, donde la especificidad de la RH nunca superó el 90%²³⁷. Esto se hace aún más sorprendente cuando hay tener en cuenta que algunas de las integraciones consideradas ilegítimas en este modelo podrían deberse a integraciones en el locus *WAS* endógeno, mediante integración del DNA donador a través de aEJ y/o mediante RH, utilizando como brazo de homología alternativo la región de 236pb del Δ LTR 3' del DNA donador (presente a 2kb' del brazo de homología correcto). Ambos procesos darían lugar a células dobles positivas (eGFP⁺dsRED⁺) que son contabilizadas como integraciones ilegítimas. Aunque esto no es un problema para poder comparar diferentes estrategias para

mejorar la eficacia y seguridad de RH, si limita la cuantificación absoluta de la RH y las inserciones ilegítimas. Es curioso que, a pesar de la sobrestimación de las inserciones del donador fuera de sitio, la especificidad detectada sea superior al 99%. Nuestra hipótesis es que esto podría deberse al alto número de dianas (13) encontradas en este modelo, lo que favorecería la RH en el sitio diana, frente a la integración del donador en sitios ilegítimos. Otro factor podría ser el tamaño de los brazos de homología (hasta 3kb en el brazo 5').

Por otro lado, el modelo K562 SE- δ GF-I₁-P nos permitió determinar que la inclusión de la diana de corte en el DNA donador mejoraba sustancialmente la eficacia de RH sin afectar negativamente a la especificidad. Esto concuerda con aproximaciones propuestas por otros grupos, donde la inclusión de la diana en el donador favorece la RH^{248,249}, así como la integración del DNA donador en la diana independiente de RH (HITI)^{254,255}.

Dadas las limitaciones de este sistema, decidimos diseñar otro modelo celular reportero que nos permitiese evaluar de forma más precisa (o complementaria) las estrategias de restauración génica. Este segundo modelo, denominado K562 SEWAS84, contenía una única copia del cassette de expresión SE intacto, precedido de un fragmento de 84pb del primer intrón de *WAS*. La eficacia de RH en este modelo se determinaría de forma diferente. En este caso, la RH silencia eGFP y/o enciende dsRED, en ambos casos se estaría produciendo RH. En este modelo, de nuevo, la doble expresión de eGFP y dsRED indicaría integración ilegítima. Este modelo sirvió para constatar que los donadores que incluyen el sitio diana para corte de CRISPR/Cas9 logran mayores eficacias de RH sin afectar la especificidad relativa.

Por otro lado, aunque este modelo no nos permitía determinar los niveles exactos de integraciones inespecíficas (en determinados casos detectamos silenciamiento de

eGFP en ausencia de donador) si podíamos establecer un rango de valores entre los que se situarían los niveles exactos de integraciones inespecíficas de cada donador. Para ello estimamos, por un lado, que todas las células negativas para eGFP y dsRED son producto de RH (bien utilizando los dos brazos del donador o solo el brazo conteniendo eGFP mutado), lo que nos daba un valor hipotético de la especificidad claramente superior a lo real. Y, por otro, que en ningún caso las células negativas para eGFP y dsRED son producto de RH, lo que nos permitía establecer el límite superior de inespecificidad, por debajo del cual estarían nuestros donadores. Esto nos permitió estimar que en este modelo las integraciones ilegítimas estaban por encima del 1,8% en el caso de SdR*eGFP, y del 13% en el caso de SdR*eGFPWAS84. Este valor es mucho más cercano al 20% de integraciones inespecífica que detectamos en una publicación previa²³⁷.

Eficiencia y especificidad de sistemas CRISPR/CAS9 para modificar mediante RH el locus *WAS* en HSCs

Nuestro tercer objetivo era analizar la eficiencia y seguridad de los sistemas CRISPR específicos para *WAS* para modificar el fenotipo normal de HSCs mediante RH.

Desde un punto de vista de la TG, las HSCs representan una de las dianas terapéuticas más relevantes, dada su capacidad para repoblar el sistema hematopoyético. La modificación genética de estas células a partir de vectores virales se ha empleado exitosamente para el tratamiento de inmunodeficiencias, como WAS y ADA-SCID, y de hemoglobinopatías como la anemia falciforme^{164,204,256}. Sin embargo, el carácter integrativo de los vectores virales empleados en este tipo de estrategias conlleva riesgo de mutagénesis insertacional, a pesar de los elevados estándares de bioseguridad de los vectores actuales. Otro factor a tener en cuenta es que la inserción aleatoria de este tipo de vectores en el genoma de las células diana provocará una heterogeneidad en la expresión del transgén, lo que podría afectar la funcionalidad de las HSCs. Frente a estas estrategias, la modificación de HSCs autólogas mediante EG representa una alternativa más prometedora para el tratamiento de este tipo de desórdenes, ya que posibilita la corrección de la mutación causante de la enfermedad de forma más fisiológica. Sin embargo, aunque el uso de endonucleasas específicas ha supuesto un impulso en la capacidad de eliminar o corregir en HSCs mutaciones asociadas a determinadas patologías, esta aproximación aún debe superar una serie de limitaciones para avanzar en su aplicación clínica. Entre los principales retos que deben abordarse, destacan: 1- lo adversas que se muestran las HSCs a su modificación, siendo las bajas eficiencias de EG alcanzadas en HSCs un problema al que se han enfrentado numerosos grupos, 2- la entrega *ex vivo* e *in vivo* de las herramientas de EG, que de forma eficiente y transitoria deben actuar en HSCs, 3- la capacidad de lograr altos niveles de repoblación de las HSCs *in vivo*, y 4- la identificación y reducción de los efectos *off target*²⁵⁷.

El trasplante de médula ósea sigue siendo el principal tratamiento para WAS, y ha progresado en términos de acondicionamiento, y selección de donantes, alcanzando actualmente una tasa de supervivencia superior al 80%^{199,202,258,259}. Sin embargo, el riesgo de sufrir complicaciones y los efectos secundarios que suceden al trasplante en pacientes con WAS, así como la imposibilidad de llevarlo a cabo en pacientes que no cuentan con donantes compatibles, han contribuido a la búsqueda de terapias alternativas. En este sentido, la obtención de HSCs autólogas expresando la proteína WASP, empleando LVs, ha demostrado en varios ensayos clínicos claros beneficios terapéuticos en pacientes con WAS^{203,204}. A pesar de ello, esta estrategia no ha logrado revertir la microtrombocitopenia^{204,260}, cuya persistencia también se ha observado después de trasplante alogénico de HSCs, asociada en este caso a un bajo número de células mieloides derivadas del donante²⁰⁰.

Nuestra hipótesis es que una alternativa terapéutica para WAS, que podría superar las limitaciones de la TG basada en el uso de vectores integrativos, sería la corrección mediante EG de la mutación causante de WAS en HSCs autólogas, lo que permitiría una expresión más fisiológica de WASP en pacientes. Esta corrección podría conseguirse mediante diferentes vías:

1. Restauración del marco de lectura mediada por NHEJ. Esta vía de reparación del DNA es más eficiente que la RH en HSCs, y permitiría la restauración del marco de lectura abierto en aproximadamente un tercio de los indels generados, tras crear un DSB en el sitio de la mutación²⁶¹. Otra alternativa sería crear dos DSBs adyacentes, que flanqueen la mutación manteniendo el marco de lectura y eliminando aminoácidos irrelevantes para la función de la proteína²⁶².
2. Corrección de la mutación mediante RH. Esta sería la estrategia ideal, ya que favorecería la expresión génica fisiológica, conservando las regiones

reguladoras endógenas²⁵⁷ (particularmente interesante en el caso de *WAS*, dada la complejidad de la regulación de su expresión²⁶³⁻²⁶⁵). Por otro lado, frente a la estrategia anterior, la RH garantiza la fidelidad en la reparación.

3. Inserción sitio-específica del cDNA de *WAS*. Esta estrategia permitiría dirigir la inserción del cassette para el cDNA de *WAS* en un puerto seguro, como el locus *AAVS1*, en que la inserción no altere la biología de la célula diana. La ventaja de esta estrategia es que, una vez diseñadas las nucleasas específicas idóneas frente al puerto seguro, estas podrían usarse para el tratamiento de diferentes trastornos genéticos, variando el DNA donador²⁵⁷.

No obstante, para que el traslado de esta alternativa a clínica sea viable, es necesario superar las barreras anteriormente mencionadas. Ante esta necesidad, nos propusimos abordar la mejora de la eficiencia y la seguridad de las herramientas de edición génica específicas para el locus *WAS* como principal objetivo de esta tesis, tal y como se ha intentado exponer.

Inicialmente apostamos por los IDLVs como un posible método de entrega del sistema CRISPR/Cas9, sobre la hipótesis de que su optimización permitiría alcanzar los niveles de expresión de Cas9 y gRNA necesarios, y suficientes, para editar HSCs de forma eficaz y segura. Sin embargo, no fuimos capaces de observar disrupción del locus *WAS* en nuestras células diana, células T y HSCs, usando nuestros IDLVs mejorados. Otros grupos también han tenido dificultades para hacer llegar NE a estas células, lo que se debería a una expresión insuficiente²⁶⁶. No obstante, los IDLVs aún podrían emplearse para la entrega de DNA donadores en estas células, tal y como demostraron Genovese et al.²⁶⁷.

En nuestras manos, el único método que finalmente permitió altas eficiencias de disrupción del locus *WAS* en HSCs y células T, fue la entrega del sistema CRISPR/Cas9 como RNP. Esto nos permitió combinar la nucleofección de RNP con la de DNA donadores en forma de plásmido, o en forma de productos de PCR para abordar la RH en HSCs. Solo en el segundo caso observamos la integración del donador en el locus *WAS* una pequeña fracción de HSCs (0,6% de la población), acompañada de una reducción en los niveles de expresión de la proteína WASP. Sin embargo, no pudimos establecer en qué proporción el DNA donador se había integrado en el locus *WAS*, y en qué proporción en sitios ilegítimos debido a débil señal obtenida en la tinción de WASP.

Aunque nuestras eficiencias de RH en HSCs fueron bastante bajas, creemos que la adaptación de nuestras herramientas a los nuevos protocolos, combinando RNP con rAAV6 para la entrega de donadores²⁶⁸, nos permitirá lograr las eficacias suficientes como para abordar estudios preclínicos en modelos animales.

CONCLUSIONES

1. El método de entrega del sistema CRISPR no tiene grandes efectos en su especificidad pero si en su eficacia, que es, además, dependiente del tipo celular.
2. La inclusión del elemento IS2 en IDLVs mejora su actividad transcripcional pero reduce la eficacia de formación de episomas, de forma que el efecto final dependerá del balance entre los efectos negativos y positivos en cada tipo celular.
3. El incremento de actividad transcripcional de los episomas IS2-IDLVs no se debe al bloqueo del silenciamiento por HDAC, ni al incremento de formas 1-LTR, sino que reposiciona los episomas en regiones transcripcionalmente más activas del núcleo.
4. En células T y HSCs, el IS2 no mejora la eficacia de los IDLVs.
5. Se han generado dos modelos que permiten analizar eficiencia y seguridad de reparación mediante RH de sistemas CRISPR/Cas9 específicos para *WAS*. Estos modelos nos han permitido concluir que:
 - a. Las integraciones del DNA donador en sitios indeseados dependen del diseño del DNA donador, así como del número de dianas en el genoma de la célula.
 - b. Los DNA donadores que incluyen la diana de corte mejoran la eficacia de RH, sin incrementar la inespecificidad.
6. La nucleofección de complejos ribonucleoproteicos (gRNA-Cas9) es el método de elección para la EG de células T y HSCs. Combinándolo con donadores como productos de PCR, permite editar el locus *WAS* en HSCs.

BIBLIOGRAFÍA

1. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & David, R. HHS Public Access. **533**, 420–424 (2016).
2. Gaudelli, N. M. *et al.* Programmable base editing of T to G C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).
3. Amabile, A. *et al.* Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell* **167**, 219–232.e14 (2016).
4. Saunderson, E. A. *et al.* Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. *Nat. Commun.* **8**, 1–14 (2017).
5. Liao, H. K. *et al.* In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation. *Cell* **171**, 1495–1507.e15 (2017).
6. Rinaldi, C. & Wood, M. J. A. Antisense oligonucleotides: The next frontier for treatment of neurological disorders. *Nat. Rev. Neurol.* **14**, 9–22 (2018).
7. Mullard, A. Pioneering antisense drug heads into pivotal trials for Huntington disease. *Nat. Rev. Drug Discov.* **18**, 161–163 (2019).
8. Felsenfeld, G.; Davies, D. R.; Rich, A. I i 0.01. *J. Am. Chem. Soc.* **79**, 2023–2024 (1957).
9. Morgan, A. R. & Wells, R. D. Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. *J. Mol. Biol.* **37**, 63–80 (1968).
10. Montazersaheb, S., Hejazi, M. S. & Nozad Charoudeh, H. Potential of Peptide Nucleic Acids in Future Therapeutic Applications. *Adv. Pharm. Bull.* **8**, 551–563 (2018).
11. Malik, S., Oyaghire, S. & Bahal, R. Applications of PNA-laden nanoparticles for hematological disorders. *Cell. Mol. Life Sci.* (2018). doi:10.1007/s00018-018-2979-5
12. Ricciardi, A. S., Quijano, E., Putman, R., Saltzman, W. M. & Glazer, P. M. Peptide nucleic acids as a tool for site-specific gene editing. *Molecules* **23**, 1–15 (2018).
13. Corradini, R. Special Issue: Molecular Properties and the Applications of Peptide Nucleic Acids. *Molecules* **23**, 1977 (2018).
14. Chin, J. Y., Reza, F. & Glazer, P. M. Triplex-forming peptide nucleic acids induce heritable elevations in gamma-globin expression in hematopoietic progenitor cells. *Mol. Ther.* **21**, 580–587 (2013).
15. Nielsen, P., Egholm, M., Berg, R. & Buchardt, O. *Peptide Nucleic Acids (PNA). Potential Antisense and Antigene Agents. Anti-cancer drug design* **8**, (1993).
16. Nielsen, P. E. Applications of peptide nucleic acids. *Curr. Opin. Biotechnol.* **10**, 71–75 (1999).
17. Nielsen, P. E. Peptide Nucleic Acids and the Origin of Life. *Chem. Biodivers.* **4**, 1996–2002 (2007).
18. Nielsen, P. E. Peptide Nucleic Acids: On the Road to New Gene Therapeutic Drugs. *Pharmacol. Toxicol.* **86**, 3–7 (2008).
19. Nielsen, P. E. Peptide Nucleic Acids (PNA) in Chemical Biology and Drug Discovery. *Chem. Biodivers.* **7**, 786–804 (2010).
20. Smith, L. J. *et al.* Stem cell-derived clade F AAVs mediate high-efficiency homologous recombination-based genome editing. *Proc. Natl. Acad. Sci.* **115**, E7379–E7388 (2018).
21. François, J.-C., Saison-Behmoaras, T. & Hélène, C. Sequence-specific recognition of the major groove of DNA by oligodeoxynucleotides via triple helix formation. Footprinting studies. *Nucleic Acids Res.* **16**, 11431–11440 (1988).
22. P.P., C. *et al.* Targeted correction of an episomal gene in mammalian cells by a short DNA fragment tethered to a triplex-forming oligonucleotide. *J. Biol. Chem.* **274**, 11541 (1999).
23. Datta, H. J., Chan, P. P., Vasquez, K. M., Gupta, R. C. & Glazer, P. M. Triplex-induced Recombination in Human Cell-free Extracts. *J. Biol. Chem.* **276**, 18018–18023 (2001).

24. Ricciardi, A. S. *et al.* In utero nanoparticle delivery for site-specific genome editing. *Nat. Commun.* **9**, 1–11 (2018).
25. McNeer, N. A. *et al.* Systemic delivery of triplex-forming PNA and donor DNA by nanoparticles mediates site-specific genome editing of human hematopoietic cells in vivo. *Gene Ther.* **20**, 658–669 (2013).
26. McNeer, N. A. that deliver triplex-forming peptide nucleic acid molecules correct F. C. in airway epithelium *et al.* Nanoparticles that deliver triplex-forming peptide nucleic acid molecules correct F508del CFTR in airway epithelium. *Nat. Commun.* **6**, 1–11 (2015).
27. Bahal, R., McNeer, N. A., Ly, D. H., Saltzman, W. M. & Glazer, P. M. Nanoparticle for delivery of antisense γ PNA oligomers targeting CCR5. *Artif. DNA PNA XNA* **4**, 43–53 (2013).
28. Schleifman, E. B. *et al.* Targeted disruption of the CCR5 gene in human hematopoietic stem cells stimulated by peptide nucleic acids. *Chem. Biol.* **18**, 1189–1198 (2011).
29. Vasileva, A., Linden, R. M. & Jessberger, R. Homologous recombination is required for AAV-mediated gene targeting. *Nucleic Acids Res.* **34**, 3345–3360 (2006).
30. Barzel, A. *et al.* Promoterless gene targeting without nucleases ameliorates haemophilia B in mice. *Nature* **517**, 360–364 (2015).
31. Charlesworth, C. T. *et al.* Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254 (2019).
32. Porteus, M. H. A New Class of Medicines through DNA Editing. *N. Engl. J. Med.* **380**, 947–959 (2019).
33. Jasin, M. Introduction of Double-Strand Breaks into the Genome of Mouse Cells by Expression of a Rare-Cutting Endonuclease. **14**, 8096–8106 (1994).
34. Choulika, A., Perrin, A., Dujon, B. & Nicolas, J. F. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **15**, 1968–73 (1995).
35. Gimble, F. S. Invasion of a multitude of genetic niches by mobile endonuclease genes. *FEMS Microbiol. Lett.* **185**, 99–107 (2000).
36. Dalgaard, J. Statistical modeling and analysis of the LAGLIDADG family of site-specific endonucleases and identification of an intein that encodes a site-specific endonuclease of the HNH family. *Nucleic Acids Res.* **25**, 4626–4638 (2002).
37. Stoddard, B. L. Homing endonucleases from mobile group I introns: Discovery to genome engineering. *Mob. DNA* **5**, 1–15 (2014).
38. Duchateau, P. *et al.* Meganucleases and Other Tools for Targeted Genome Engineering: Perspectives and Challenges for Gene Therapy. *Curr. Gene Ther.* **11**, 11–27 (2011).
39. Stoddard, B. L. Homing endonucleases from mobile group I introns: Discovery to genome engineering. *Mob. DNA* **5**, 1–15 (2014).
40. Seligman, L. M. Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res.* **30**, 3870–3879 (2002).
41. Sussman, D. *et al.* Isolation and characterization of new homing endonuclease specificities at individual target site positions. *J. Mol. Biol.* **342**, 31–41 (2004).
42. Rosen, L. E. *et al.* Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Res.* **34**, 4791–4800 (2006).
43. Doyon, J. B., Pattanayak, V., Meyer, C. B. & Liu, D. R. Directed evolution and substrate specificity profile of homing endonuclease I-SceI. *J. Am. Chem. Soc.* **128**, 2477–2484 (2006).
44. Chen, Z., Wen, F., Sun, N. & Zhao, H. Directed evolution of homing endonuclease I-SceI with altered sequence specificity. *Protein Eng. Des. Sel.* **22**, 249–256 (2009).
45. Joshi, R. *et al.* recognition site specificity. **405**, 185–200 (2012).

46. Chevalier, B. S. *et al.* Design, activity, and structure of a highly specific artificial endonuclease. *Mol. Cell* **10**, 895–905 (2002).
47. Epinat, J. C. *et al.* A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res.* **31**, 2952–2962 (2003).
48. Silva, G. H., Belfort, M., Wende, W. & Pingoud, A. From Monomeric to Homodimeric Endonucleases and Back: Engineering Novel Specificity of LAGLIDADG Enzymes. *J. Mol. Biol.* **361**, 744–754 (2006).
49. Baxter, S. *et al.* Engineering domain fusion chimeras from I-Onul family LAGLIDADG homing endonucleases. *Nucleic Acids Res.* **40**, 7985–8000 (2012).
50. Arnould, S. *et al.* Engineered I-Crel Derivatives Cleaving Sequences from the Human XPC Gene can Induce Highly Efficient Gene Correction in Mammalian Cells. *J. Mol. Biol.* **371**, 49–65 (2007).
51. Grizot, S. *et al.* Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. *Nucleic Acids Res.* **37**, 5405–5419 (2009).
52. Press, N. B. & li, C. New restriction endonucleases from *Flavobacterium okeanokoites* (FokI) and *Micrococcus luteus* Hiroyuki Sugisaki and Susumu Kanazawa. **16**, 73–78 (1981).
53. Chandrasegaran, S. & Carroll, D. Origins of Programmable Nucleases for Genome Engineering. *J. Mol. Biol.* **428**, 963–989 (2016).
54. Kim, Y.-G. & Chandrasegaran, S. Chimeric restriction endonuclease (*Flavobacterium okeanokoites*/*Escherichia coli*/hybrid restriction endonuclease/protein engineering/recognition and cleavage domains). *Biochemistry* **91**, 883–887 (1994).
55. Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1156–60 (1996).
56. Pavletich, N. P. & Pabo, C. O. Zinc Finger-DNA Recognition: Crystal Structure of a Zif268-DNA Complex at $\overset{\circ}{\text{A}}$. *Science (80-.)* **252**, 809–817 (1991).
57. THIESEN, H.-J. & BACH, C. DNA Recognition of C2H2 Zinc-Finger Proteins. *Ann. N. Y. Acad. Sci.* **684**, 246–249 (1993).
58. Choo, Y. Correction: Toward a Code for the Interactions of Zinc Fingers with DNA: Selection of Randomized Fingers Displayed on Phage. *Proc. Natl. Acad. Sci.* **92**, 646b – 646 (2006).
59. Razin, S. V., Borunova, V. V., Maksimenko, O. G. & Kantidze, O. L. Cys2His2 zinc finger protein family: Classification, functions, and major members. *Biochem.* **77**, 217–226 (2012).
60. Dogan, B. A DNA recognition code for probing the in vivo functions of zinc finger transcription factors at domain resolution. (2019).
61. Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636–646 (2010).
62. Bitinaite, J., Wah, D. A., Aggarwal, A. K. & Schildkraut, I. FokI dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci.* **95**, 10570 LP – 10575 (1998).
63. Smith, J. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* **28**, 3361–3369 (2000).
64. Vanamee, É. S., Santagata, S. & Aggarwal, A. K. FokI requires two specific DNA sites for cleavage. *J. Mol. Biol.* **309**, 69–78 (2001).
65. Miller, J. C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* **25**, 778 (2007).
66. Szczepek, M. *et al.* Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.* **25**, 786–793 (2007).

67. Paschon, D. E. *et al.* Diversifying the structure of zinc finger nucleases for high-precision genome editing. *Nat. Commun.* **10**, 1133 (2019).
68. Bibikova, M. *et al.* Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Mol. Cell. Biol.* **21**, 289 LP – 297 (2001).
69. Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. Enhancing Gene Targeting with Designed Zinc Finger Nucleases. *Science (80-.)*. **300**, 764 LP – 764 (2003).
70. Porteus, M. H. & Baltimore, D. Chimeric Nucleases Stimulate Gene Targeting in Human Cells. *Science (80-.)*. **300**, 763 LP – 763 (2003).
71. Urnov, F. D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).
72. Geurts, A. M. *et al.* Knockout Rats via Embryo Microinjection of Zinc-Finger Nucleases. *Science (80-.)*. **325**, 433 LP – 433 (2009).
73. Wood, A. J. *et al.* Targeted Genome Editing Across Species Using ZFNs and TALENs. *Science (80-.)*. **333**, 307 LP – 307 (2011).
74. Tebas, P. *et al.* Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV . *N. Engl. J. Med.* **370**, 901–910 (2014).
75. Holmes, M. C. *et al.* A Potential Therapy for Beta-Thalassemia (ST-400) and Sickle Cell Disease (BIVV003). *Blood* **130**, 2066 LP – 2066 (2017).
76. Schmidt, C. *et al.* Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics* **186**, 757–761 (2010).
77. Jiang, W. Z. *et al.* TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* **39**, 359–372 (2010).
78. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, 1–11 (2011).
79. Mahfouz, M. M. *et al.* De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc. Natl. Acad. Sci.* **108**, 2623–2628 (2011).
80. Sontheimer, E. J. & Marraffini, L. A. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* **11**, 181–190 (2010).
81. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci.* **109**, E2579–E2586 (2012).
82. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. Nucleotide Sequence of the *iap* Gene , Responsible for Alkaline Phosphatase Isozyme Conversion in *Escherichia coli* , and Identification of the Gene Product. 5429–5433 (1987).
83. Mojica, F. J. M., Díez-Villaseñor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* **36**, 244–246 (2000).
84. Jansen, R., Embden, J. D. A. Van, Gastra, W. & Schouls, L. M. Jansen_Identification of genes that are associated with DNA repeats in prokaryotes.pdf. **43**, 1565–1575 (2002).
85. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182 (2005).
86. Haft, D. H., Selengut, J., Mongodin, E. F. & Nelson, K. E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **1**, 0474–0483 (2005).
87. Makarova, K. S. *et al.* Evolution and classification of the CRISPR–Cas systems. *Nat. Rev. Microbiol.* **9**, 467 (2011).

88. Cong, L. *et al.* Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., ... Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* (New York, N.Y.). **339**, 819–823 (2013).
89. Sapranaukas, R. *et al.* The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* **39**, 9275–9282 (2011).
90. Thompson, L. H. & Schild, D. Recombinational DNA repair and human disease. *Mutat. Res. Mol. Mech. Mutagen.* **509**, 49–78 (2002).
91. Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **18**, 495–506 (2017).
92. Her, J. & Bunting, S. F. How cells ensure correct repair of DNA double-strand breaks. *J. Biol. Chem.* **293**, 10502–10511 (2018).
93. Takata, M. *et al.* Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* **17**, 5497–5508 (1998).
94. Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* **7**, 2902–2906 (2008).
95. Takashima, Y. *et al.* Dependence of DNA double strand break repair pathways on cell cycle phase in human lymphoblastoid cells. *Environ. Mol. Mutagen.* **50**, 815–822 (2009).
96. Bétermier, M., Bertrand, P. & Lopez, B. S. Is Non-Homologous End-Joining Really an Inherently Error-Prone Process? *PLoS Genetics* **10**, (2014).
97. Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **18**, 495 (2017).
98. Fattah, F. *et al.* Ku Regulates the Non-Homologous End Joining Pathway Choice of DNA Double-Strand Break Repair in Human Somatic Cells. *PLoS Genet.* **6**, e1000855 (2010).
99. Davis, A. J. & Chen, D. J. DNA double strand break repair via non-homologous end-joining. *Transl. Cancer Res.* **2**, 130–143 (2013).
100. Chu, G., DeFazio, L. G., Stansel, R. M. & Griffith, J. D. Synapsis of DNA ends by DNA-dependent protein kinase. *EMBO J.* **21**, 3192–3200 (2002).
101. Chang, H. H. Y. & Lieber, M. R. Structure-Specific nuclease activities of Artemis and the Artemis: DNA-PKcs complex. *Nucleic Acids Res.* **44**, 4991–4997 (2016).
102. Weterings, E. & Chen, D. J. The endless tale of non-homologous end-joining. *Cell Res.* **18**, 114–24 (2008).
103. Reid, D. A. *et al.* Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair. *Proc. Natl. Acad. Sci.* **112**, E2575–E2584 (2015).
104. Williams, G. J. *et al.* Structural insights into NHEJ: Building up an integrated picture of the dynamic DSB repair super complex, one component and interaction at a time. *DNA Repair (Amst)*. **17**, 110–120 (2014).
105. Ousterout, D. G. *et al.* Reading frame correction by targeted genome editing restores dystrophin expression in cells from duchenne muscular dystrophy patients. *Mol. Ther.* **21**, 1718–1726 (2013).
106. Bassel-Duby, R. *et al.* Correction of diverse muscular dystrophy mutations in human engineered heart muscle by single-site genome editing. *Sci. Adv.* **4**, eaap9004 (2018).
107. Perez, E. E. *et al.* Establishment of HIV-1 resistance in CD4 + T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816 (2008).

108. Ren, J. *et al.* Multiplex Genome Editing to Generate Universal CAR T Cells Resistant to PD1 Inhibition. *Clin. Cancer Res.* **23**, 2255–2266 (2017).
109. Chakrabarti, A. M. *et al.* Target-Specific Precision of CRISPR-Mediated Genome Editing. *Mol. Cell* **73**, 699–713.e6 (2019).
110. Rothkamm, K., Krüger, I., Thompson, L. H. & Löbrich, M. Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol. Cell. Biol.* **23**, 5706–5715 (2003).
111. Branzei, D. & Foiani, M. Regulation of DNA repair throughout the cell cycle. *Nat. Rev. Mol. Cell Biol.* **9**, 297 (2008).
112. San Filippo, J., Sung, P. & Klein, H. Mechanism of Eukaryotic Homologous Recombination. *Annu. Rev. Biochem.* **77**, 229–257 (2008).
113. Sung, P. & Klein, H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat. Rev. Mol. Cell Biol.* **7**, 739–750 (2006).
114. Carson, C. T. *et al.* The Mre11 complex is required for ATM activation and the G2/M checkpoint. *EMBO J.* **22**, 6610–6620 (2003).
115. West, S. C. Molecular views of recombination proteins and their control. *Nat Rev Mol Cell Biol* **4**, 435–445 (2003).
116. Wyman, C. & Kanaar, R. DNA Double-Strand Break Repair: All's Well that Ends Well. *Annu. Rev. Genet.* **40**, 363–383 (2006).
117. Alani, E., Thresher, R., Griffith, J. D. & Kolodner, R. D. Characterization of DNA-binding and strand-exchange stimulation properties of γ -RPA, a yeast single-strand-DNA-binding protein. *J. Mol. Biol.* **227**, 54–71 (1992).
118. Liu, Y., Masson, J., Shah, R., Regan, P. O. & West, S. C. RAD51C Is Required for Holliday. *Spring* **303**, 243–246 (2004).
119. Smith, C. *et al.* Whole-Genome Sequencing Analysis Reveals High Specificity of CRISPR/Cas9 and TALEN-Based Genome Editing in Human iPSCs. *Cell Stem Cell* **15**, 12–13 (2014).
120. Veres, A. *et al.* Low Incidence of Off-Target Mutations in Individual CRISPR-Cas9 and TALEN Targeted Human Stem Cell Clones Detected by Whole-Genome Sequencing. *Cell Stem Cell* **15**, 254 (2014).
121. Yang, L. *et al.* Targeted and genome-wide sequencing reveal single nucleotide variations impacting specificity of Cas9 in human stem cells. *Nat. Commun.* **5**, 1–6 (2014).
122. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168 (2014).
123. Hsiao, T. *et al.* Inference of CRISPR Edits from Sanger Trace Data. *bioRxiv* 251082 (2019). doi:10.1101/251082
124. Sentmanat, M. F., Peters, S. T., Florian, C. P., Connelly, J. P. & Pruett-Miller, S. M. A Survey of Validation Strategies for CRISPR-Cas9 Editing. *Sci. Rep.* **8**, 1–8 (2018).
125. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
126. Gabriel, R. *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* (2011). doi:10.1038/nbt.1948
127. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187 (2014).
128. Frock, R. L. *et al.* Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* **33**, 179–188 (2015).
129. Crosetto, N. *et al.* Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* **10**, 361–365 (2013).
130. Yan, W. X. *et al.* BLISS is a versatile and quantitative method for genome-wide profiling of

- DNA double-strand breaks. *Nat. Commun.* **8**, 1–9 (2017).
131. Kuscü, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* **32**, 677–683 (2014).
 132. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* (2015). doi:10.1038/nbt.3117
 133. Friedland, A. E. *et al.* Characterization of *Staphylococcus aureus* Cas9: A smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. *Genome Biol.* (2015). doi:10.1186/s13059-015-0817-8
 134. Amrani, N. *et al.* NmeCas9 is an intrinsically high-fidelity genome-editing platform. *Genome Biol.* **19**, 214 (2018).
 135. Park, P. J. ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* **10**, 669–680 (2009).
 136. Kuscü, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* **32**, 677 (2014).
 137. Cencic, R. *et al.* Protospacer adjacent motif (PAM)-distal sequences engage CRISPR/Cas9 DNA target cleavage. *PLoS One* **9**, 1–13 (2014).
 138. Martin, F., Sánchez-Hernández, S., Gutiérrez-Guerrero, A., Pinedo-Gomez, J. & Benabdellah, K. Biased and unbiased methods for the detection of off-target cleavage by CRISPR/Cas9: An overview. *Int. J. Mol. Sci.* **17**, (2016).
 139. Campos, S. K. & Barry, M. A. Current advances and future challenges in Adenoviral vector biology and targeting. *Curr. Gene Ther.* **7**, 189–204 (2007).
 140. Schagen, F. H. E., Ossevoort, M., Toes, R. E. M. & Hoeben, R. C. Immune responses against adenoviral vectors and their transgene products: A review of strategies for evasion. *Crit. Rev. Oncol. Hematol.* **50**, 51–70 (2004).
 141. Wold, W. S. M. & Toth, K. Adenovirus vectors for gene therapy, vaccination and cancer gene therapy. *Curr. Gene Ther.* **13**, 421–33 (2013).
 142. Li, C. *et al.* Inhibition of HIV-1 infection of primary CD4+ T-cells by gene editing of CCR5 using adenovirus-delivered CRISPR/Cas9. *J. Gen. Virol.* **96**, 2381–2393 (2015).
 143. Maggio, I. *et al.* Selection-free gene repair after adenoviral vector transduction of designer nucleases: Rescue of dystrophin synthesis in DMD muscle cell populations. *Nucleic Acids Res.* **44**, 1449–1470 (2016).
 144. Ogradowski, P. *et al.* In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* **516**, 423–427 (2014).
 145. Wang, D. *et al.* Adenovirus-Mediated Somatic Genome Editing of Pten by CRISPR/Cas9 in Mouse Liver in Spite of Cas9-Specific Immune Responses. *Hum. Gene Ther.* **26**, 432–442 (2015).
 146. Cheng, R. *et al.* Efficient gene editing in adult mouse livers via adenoviral delivery of CRISPR/Cas9. *FEBS Lett.* **588**, 3954–3958 (2014).
 147. Xu, L. *et al.* CRISPR-mediated genome editing restores dystrophin expression and function in mdx mice. *Mol. Ther.* **24**, 564–569 (2016).
 148. Stephens, C. J. *et al.* Long-term correction of hemophilia B using adenoviral delivery of CRISPR/Cas9. *J. Control. Release* **298**, 128–141 (2019).
 149. Rose, J. A., Berns, K. I., Hoggan, M. D. & Koczot, F. J. Evidence for a Single-Stranded Adenovirus-Associated Virus Genome: Formation of a Dna Density Hybrid on Release of Viral Dna. *Proc. Natl. Acad. Sci.* **64**, 863–869 (2006).
 150. Koczot, F. J., Carter, B. J., Garon, C. F. & Rose, J. A. Self-Complementarity of Terminal

- Sequences Within Plus or Minus Strands of Adenovirus-Associated Virus DNA. *Proc. Natl. Acad. Sci.* **70**, 215–219 (2006).
151. Athanasopoulos, T., Munye, M. M. & Yáñez-Muñoz, R. J. Nonintegrating Gene Therapy Vectors. *Hematol. Oncol. Clin. North Am.* **31**, 753–770 (2017).
 152. Lusby, E., Fife, K. H. & Berns, K. I. Nucleotide Sequence of the Inverted Terminal Repetition in Adeno-Associated Virus DNA. *DNA Seq.* **34**, 402–409 (1980).
 153. Colella, P., Ronzitti, G. & Mingozzi, F. Emerging Issues in AAV-Mediated In Vivo Gene Therapy. *Mol. Ther. - Methods Clin. Dev.* **8**, 87–104 (2018).
 154. Kotterman, M. A. & Schaffer, D. V. Engineering adeno-associated viruses for clinical gene therapy. *Nat. Rev. Genet.* **15**, 445–451 (2014).
 155. Blaese, R. M. *et al.* T Lymphocyte-Directed Gene Therapy for ADA^{scid};-</sup> SCID: Initial Trial Results After 4 Years. *Science (80-)*, **270**, 475 LP – 480 (1995).
 156. Bordignon, C. *et al.* Gene Therapy in Peripheral Blood Lymphocytes and Bone Marrow for ADA-. *Source Sci. New Ser. Sci. J. Mol. Biol. Plant Mol. Biol. Le Guen a!. Mol. Gen. Genet* **270**, 470–475 (1995).
 157. Kohn, D. B. *et al.* Engraftment of gene-modified umbilical cord blood cells in Kohn, D. B., Weinberg, K. I., Nolta, J. A., Heiss, L. N., Lenarsky, C., Crooks, G. M., ··· el-Khoureiy, A. (1995). Engraftment of gene-modified umbilical cord blood cells in neonates with adenosine . *Nat. Med.* **1**, 1017–23 (1995).
 158. Kohn, D. B. *et al.* T lymphocytes with a normal ADA gene accumulate after transplantation of transduced autologous umbilical cord blood CD34+ cells in ADA-deficient SCID neonates. *Nat. Med.* **4**, 775–780 (1998).
 159. Fischer, A., Hacein-Bey-Abina, S. & Cavazzana-Calvo, M. 20 years of gene therapy for SCID. *Nat. Immunol.* **11**, 457–460 (2010).
 160. Cavazzana-Calvo, M. *et al.* Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science (80-)*, **288**, 669–672 (2000).
 161. Alessandro Aiuti *et al.* Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science (80-)*, **296**, 5 (2002).
 162. Hacein-Bey-Abina, S. *et al.* Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N Engl J Med* **346**, 1185–1193 (2002).
 163. Gaspar, H. B. *et al.* Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet* **364**, 2181–2187 (2004).
 164. Aiuti, A. *et al.* Gene Therapy for Immunodeficiency Due to Adenosine Deaminase Deficiency. *N. Engl. J. Med.* **360**, 447–458 (2009).
 165. Wang, T. Y., Han, Z. M., Chai, Y. R. & Zhang, J. H. A mini review of MAR-binding proteins. *Mol. Biol. Rep.* **37**, 3553–3560 (2010).
 166. Hagedorn, C., Antoniou, M. N. & Lipps, H. J. Genomic cis-acting Sequences Improve Expression and Establishment of a Nonviral Vector. *Mol. Ther. Nucleic Acids* **2**, e118 (2013).
 167. Yusufzai, T. M. & Felsenfeld, G. The 5'-HS4 chicken -globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc. Natl. Acad. Sci.* **101**, 8620–8624 (2004).
 168. Zielke, K. *et al.* The Insulator Protein CTCF Binding Sites in the orf73/LANA Promoter Region of Herpesvirus Saimiri Are Involved in Conferring Episomal Stability in Latently Infected Human T Cells. *J. Virol.* (2012). doi:10.1128/jvi.06295-11
 169. Holohan, E. E. *et al.* CTCF genomic binding sites in *Drosophila* and the organisation of the bithorax complex. *PLoS Genet.* (2007). doi:10.1371/journal.pgen.0030112
 170. Mohan, M. *et al.* The *Drosophila* insulator proteins CTCF and CP190 link enhancer

- blocking to body patterning. *EMBO J.* (2007). doi:10.1038/sj.emboj.7601851
171. Herold, M., Bartkuhn, M. & Renkawitz, R. CTCF: insights into insulator function during development. *Development* **139**, 1045–57 (2012).
 172. Benabdellah, K., Gutierrez-Guerrero, A., Cobo, M., Muñoz, P. & Martín, F. A chimeric HS4-SAR insulator (IS2) that prevents silencing and enhances expression of lentiviral vectors in pluripotent stem cells. *PLoS One* **9**, (2014).
 173. Jin, Y. *et al.* Mutations of the Wiskott-Aldrich Syndrome Protein (WASP): hotspots , effect on transcription , and translation and phenotype / genotype correlation. **104**, 4010–4019 (2004).
 174. Massaad, M. J., Ramesh, N. & Geha, R. S. Wiskott-Aldrich syndrome: a comprehensive review. *Ann. N. Y. Acad. Sci.* **1285**, 26–43 (2013).
 175. Catucci, M., Castiello, M. C., Pala, F., Bosticardo, M. & Villa, A. Autoimmunity in Wiskott-Aldrich syndrome: An unsolved enigma. *Front. Immunol.* **3**, 1–14 (2012).
 176. ALDRICH, R. A., STEINEBERG, A. G. & CAMPBELL, D. C. PEDIGREE DEMONSTRATING A SEX-LINKED RECESSIVE CONDITION CHARACTERIZED BY DRAINING EARS, ECZEMATOID DERMATITIS AND BLOODY DIARRHEA. *Pediatrics* **13**, 133 LP – 139 (1954).
 177. Derry, J. M., Ochs, H. D. & Francke, U. Isolation of a novel gene mutated in Wiskott-Aldrich syndrome [published erratum appears in Cell 1994 Dec 2;79(5):following 922]. *Cell* **78**, 635–644 (1994).
 178. Ochs, H. D. Mutations of the Wiskott-Aldrich Syndrome Protein affect protein expression and dictate the clinical phenotypes. *Immunol. Res.* **44**, 84–88 (2009).
 179. Chandra, S., Bronicki, L., Nagaraj, C. B. & Zhang, K. WAS -Related Disorders Summary Diagnosis. (2019).
 180. Villa, A. *et al.* X-linked thrombocytopenia and Wiskott–Aldrich syndrome are allelic diseases with mutations in the WASP gene. *Nat. Genet.* **9**, 414–417 (1995).
 181. Devriendt, K. *et al.* Constitutively activating mutation in WASP causes X-linked severe congenital neutropenia. *Nat. Genet.* **27**, 313–317 (2001).
 182. Buchbinder, D., Nugent, D. J. & Fillipovich, A. H. Wiskott-Aldrich syndrome: Diagnosis, current management, and emerging treatments. *Appl. Clin. Genet.* **7**, 55–66 (2014).
 183. Snapper, S. B. & Rosen, F. S. THE WISKOTT-ALDRICH SYNDROME PROTEIN (WASP): Roles in Signaling and Cytoskeletal Organization. *Annu. Rev. Immunol.* **17**, 905–929 (1999).
 184. Calle, Y., Chou, H. C., Thrasher, A. J. & Jones, G. E. Wiskott-Aldrich syndrome protein and the cytoskeletal dynamics of dendritic cells. *J. Pathol.* **204**, 460–469 (2004).
 185. Thrasher, A. J. & Burns, S. O. WASP: a key immunological multitasker. *Nat. Rev. Immunol.* **10**, 182 (2010).
 186. Fearon, E. R., Kohn, D. B., Winkelstein, J. A., Vogelstein, B. & Blaese, R. M. Carrier detection in the Wiskott Aldrich syndrome. *Blood* **72**, 1735 LP – 1739 (1988).
 187. Greer, W. L. *et al.* X-Chromosome inactivation in the Wiskott-Aldrich syndrome: A marker for detection of the carrier state and identification of cell lineages expressing the gene defect. *Genomics* **4**, 60–67 (1989).
 188. Parolini, O. *et al.* X-Linked Wiskott–Aldrich Syndrome in a Girl. *N. Engl. J. Med.* **338**, 291–295 (1998).
 189. Andreu, N. *et al.* Wiskott–Aldrich syndrome in a female with skewed X-chromosome inactivation. *Blood Cells, Mol. Dis.* **31**, 332–337 (2003).
 190. Massaad, M. J., Ramesh, N. & Geha, R. S. Wiskott-Aldrich syndrome: a comprehensive review. *Ann. N. Y. Acad. Sci.* **1285**, 26–43 (2013).
 191. Pollitt, A. Y. & Insall, R. H. WASP and SCAR/WAVE proteins: the drivers of actin assembly. *J. Cell Sci.* **122**, 2575–2578 (2009).

192. Kurisu, S. & Takenawa, T. The WASP and WAVE family proteins. *Genome Biol.* **10**, 226 (2009).
193. Takenawa, T. & Suetsugu, S. The WASP–WAVE protein network: connecting the membrane to the cytoskeleton. *Nat. Rev. Mol. Cell Biol.* **8**, 37 (2007).
194. Blundell, M. P., Worth, A., Bouma, G. & Thrasher, A. J. The Wiskott-Aldrich syndrome: The actin cytoskeleton and immune cell function. *Dis. Markers* **29**, 157–175 (2010).
195. Westerberg, L. S. *et al.* WASP confers selective advantage for specific hematopoietic cell populations and serves a unique role in marginal zone B-cell homeostasis and function. *Blood* **112**, 4139–4147 (2008).
196. Litzman, J. *et al.* Intravenous immunoglobulin, splenectomy, and antibiotic prophylaxis in Wiskott-Aldrich syndrome. *Arch. Dis. Child.* **75**, 436–439 (1996).
197. N., M. *et al.* Characteristics and outcome of early-onset, severe forms of Wiskott-Aldrich syndrome. *Blood* **121**, 1510–1516 (2013).
198. Worth, A. J. J. & Thrasher, A. J. Current and emerging treatment options for Wiskott–Aldrich syndrome. *Expert Rev. Clin. Immunol.* **11**, 1015–1032 (2015).
199. Ozsahin, H. *et al.* Long-term outcome following hematopoietic stem-cell transplantation in Wiskott-Aldrich syndrome: collaborative study of the European Society for Immunodeficiencies and European Group for Blood and Marrow Transplantation. **111**, 439–445 (2008).
200. Moratto, D. *et al.* Long-term outcome and lineage-specific chimerism in 194 patients with Wiskott-Aldrich syndrome treated by hematopoietic cell transplantation in the period 1980-2009: An international collaborative study. *Blood* **118**, 1675–1684 (2011).
201. Parkman, R. *et al.* Complete Correction of the Wiskott-Aldrich Syndrome by Allogeneic Bone-Marrow Transplantation. *N. Engl. J. Med.* **298**, 921–927 (1978).
202. Shin, C. R. *et al.* Outcomes following hematopoietic cell transplantation for Wiskott–Aldrich syndrome. *Bone Marrow Transplant.* **47**, 1428 (2012).
203. Boztug, K. *et al.* Stem-Cell Gene Therapy for the Wiskott–Aldrich Syndrome. *N. Engl. J. Med.* **363**, 1918–1927 (2010).
204. Aiuti, A. *et al.* Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. *Science (80-.).* **341**, 1233151 (2013).
205. Braun, C. J. *et al.* Gene Therapy for Wiskott-Aldrich Syndrome—Long-Term Efficacy and Genotoxicity. *Sci. Transl. Med.* **6**, 227ra33 LP-227ra33 (2014).
206. Muñoz, P. *et al.* Specific marking of hESCs-derived hematopoietic lineage by WAS-promoter driven lentiviral vectors. *PLoS One* **7**, (2012).
207. Frecha, C. *et al.* Improved lentiviral vectors for Wiskott-Aldrich syndrome gene therapy mimic endogenous expression profiles throughout haematopoiesis. *Gene Ther.* **15**, 930–941 (2008).
208. Martín, F. *et al.* Lentiviral vectors transcriptionally targeted to hematopoietic cells by WASP gene proximal promoter sequences. *Gene Ther.* **12**, 715–723 (2005).
209. Cesana, D. *et al.* Uncovering and Dissecting the Genotoxicity of Self-inactivating Lentiviral Vectors In Vivo. *Mol Ther* **22**, 774–785 (2014).
210. Montes, R. *et al.* Generation and characterization of the human iPSC line PBMC1-iPS4F1 from adult peripheral blood mononuclear cells. *Stem Cell Res.* **15**, 614–617 (2015).
211. Burns, J. C., Friedmann, T., Driever, W., Burrascano, M. & Yee, J. K. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc. Natl. Acad. Sci.* **90**, 8033 LP – 8037 (1993).
212. Martín, F. *et al.* Stable Genetic Modification of Mesenchymal Stromal Cells Using Lentiviral

- Vectors. in *Viral Vectors for Gene Therapy: Methods and Protocols* (eds. Manfredsson, F. P. & Benskey, M. J.) 267–280 (Springer New York, 2019). doi:10.1007/978-1-4939-9065-8_17
213. Hendel, A. *et al.* Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* **33**, 985–989 (2015).
214. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* **25**, 402–408 (2001).
215. Ran, F. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
216. Martin Jinek *et al.* A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.)*. **337**, 816–821 (2012).
217. Wilson, I. A. *et al.* The structure of an antigenic determinant in a protein. *Cell* **37**, 767–778 (1984).
218. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
219. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).
220. Cameron, P. *et al.* Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat. Methods* **14**, 600–606 (2017).
221. Zufferey, R., Donello, J. E., Trono, D. & Hope, T. J. Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J. Virol.* **73**, 2886–92 (1999).
222. Thomas, P. & Smart, T. G. HEK293 cell line: A vehicle for the expression of recombinant proteins. *J. Pharmacol. Toxicol. Methods* **51**, 187–200 (2005).
223. Pelascini, L. P. L., Janssen, J. M. & Gonçalves, M. A. F. V. Histone Deacetylase Inhibition Activates Transgene Expression from Integration-Defective Lentiviral Vectors in Dividing and Non-Dividing Cells. *Hum. Gene Ther.* **24**, 78–96 (2012).
224. Pelascini, L. P. L. *et al.* Histone Deacetylase Inhibition Rescues Gene Knockout Levels Achieved with Integrase-Defective Lentiviral Vectors Encoding Zinc-Finger Nucleases. *Hum. Gene Ther. Methods* **24**, 399–411 (2013).
225. Ramezani, A. Performance- and safety-enhanced lentiviral vectors containing the human interferon- scaffold attachment region and the chicken γ -globin insulator. *Blood* **101**, 4717–4724 (2003).
226. Arumugam, P. I. *et al.* The 3' region of the chicken hypersensitive site-4 insulator has properties similar to its core and is required for full insulator activity. *PLoS One* **4**, (2009).
227. Urbinati, F. *et al.* Mechanism of reduction in titers from lentivirus vectors carrying large inserts in the 3' LTR. *Mol. Ther.* **17**, 1527–1536 (2009).
228. Arumugam, P. I. *et al.* Improved Human β -globin expression from self-inactivating lentiviral vectors carrying the chicken hypersensitive site-4 (cHS4) insulator element. *Mol. Ther.* **15**, 1863–1871 (2007).
229. Chung, J. H., Bell, A. C. & Felsenfeld, G. Characterization of the chicken β -globin insulator. *Proc. Natl. Acad. Sci.* **94**, 575 LP – 580 (1997).
230. Shaw, A. M. *et al.* Differences in vector-genome processing and illegitimate integration of non-integrating lentiviral vectors. *Gene Ther.* **24**, 12 (2016).
231. Ladrán, I., Tran, N., Topol, A. & Brennand, K. J. Neural stem and progenitor cells in health and disease. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **5**, 701–715 (2013).
232. Moody, S. A., Quigg, M. S. & Frankfurter, A. Development of the peripheral trigeminal system in the chick revealed by an isotype-specific anti-beta-tubulin monoclonal

- antibody. *J. Comp. Neurol.* **279**, 567–580 (1989).
233. Lee, M. K., Tuttle, J. B., Rebhun, L. I., Cleveland, D. W. & Frankfurter, A. The expression and posttranslational modification of a neuron-specific β -tubulin isotype during chick embryogenesis. *Cell Motil.* **17**, 118–132 (1990).
234. Lee, M. K., Rebhun, L. I. & Frankfurter, A. Posttranslational modification of class III beta-tubulin. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 7195–7199 (1990).
235. Easter Jr., S. S., Ross, L. S. & Frankfurter, A. Initial tract formation in the mouse brain. *J. Neurosci.* **13**, 285–299 (1993).
236. Rapti, K. *et al.* Effectiveness of gene delivery systems for pluripotent and differentiated cells. *Mol. Ther. - Methods Clin. Dev.* **2**, 14067 (2015).
237. Gutierrez-Guerrero, A. *et al.* Comparison of Zinc Finger Nucleases Versus CRISPR-Specific Nucleases for Genome Editing of the Wiskott-Aldrich Syndrome Locus. *Hum. Gene Ther.* **29**, 366–380 (2018).
238. Matz, M. V. *et al.* Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* **17**, 969–973 (1999).
239. Hinz, J. M., Laughery, M. F. & Wyrick, J. J. Nucleosomes Inhibit Cas9 Endonuclease Activity in Vitro. *Biochemistry* **54**, 7063–7066 (2015).
240. Isaac, R. S. *et al.* Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. *Elife* **5**, 1–14 (2016).
241. Horlbeck, M. A. *et al.* Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife* **5**, 1–21 (2016).
242. Chen, X. *et al.* Probing the impact of chromatin conformation on genome editing tools. *Nucleic Acids Res.* **44**, 6482–6492 (2016).
243. Daer, R. M., Cutts, J. P., Brafman, D. A. & Haynes, K. A. The Impact of Chromatin Dynamics on Cas9-Mediated Genome Editing in Human Cells. *ACS Synth. Biol.* **6**, 428–438 (2017).
244. Uusi-Mäkelä, M. I. E. *et al.* Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (*Danio rerio*). *PLoS One* **13**, 1–15 (2018).
245. Yarrington, R. M., Verma, S., Schwartz, S., Trautman, J. K. & Carroll, D. Nucleosomes inhibit target cleavage by CRISPR-Cas9 in vivo. *Proc. Natl. Acad. Sci.* **115**, 9351–9358 (2018).
246. Gisler, S. *et al.* Multiplexed Cas9 targeting reveals genomic location effects and gRNA-based staggered breaks influencing mutation efficiency. *Nat. Commun.* **10**, (2019).
247. Sánchez-Hernández, S. *et al.* The IS2 Element Improves Transcription Efficiency of Integration-Deficient Lentiviral Vector Episomes. *Mol. Ther. - Nucleic Acids* **13**, 16–28 (2018).
248. Song, F. & Stieger, K. Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks. *Mol. Ther. Nucleic Acid* **7**, 53–60 (2017).
249. Zhang, J. P. *et al.* Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* **18**, 1–18 (2017).
250. Modarai, S. R. *et al.* Efficient Delivery and Nuclear Uptake Is Not Sufficient to Detect Gene Editing in CD34+ Cells Directed by a Ribonucleoprotein Complex. *Mol. Ther. - Nucleic Acids* (2018). doi:10.1016/j.omtn.2018.01.013
251. Schumann, K. *et al.* Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc. Natl. Acad. Sci.* **112**, 10437–10442 (2015).
252. Vakulskas, C. A. *et al.* A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* (2018). doi:10.1038/s41591-018-0137-0
253. Lattanzi, A. *et al.* Optimization of CRISPR/Cas9 Delivery to Human Hematopoietic Stem and Progenitor Cells for Therapeutic Genomic Rearrangements. *Molecular Therapy*

- (2018). doi:10.1016/j.ymthe.2018.10.008
254. Suzuki, K. *et al.* In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* (2016). doi:10.1038/nature20565
255. Suzuki, K. & Izpisua Belmonte, J. C. In vivo genome editing via the HITI method as a tool for gene therapy. *Journal of Human Genetics* (2018). doi:10.1038/s10038-017-0352-4
256. Ribeil, J.-A. *et al.* Gene Therapy in a Patient with Sickle Cell Disease. *N. Engl. J. Med.* **376**, 848–855 (2017).
257. Romito, M., Rai, R., Thrasher, A. J. & Cavazza, A. Genome editing for blood disorders: state of the art and recent advances. *Emerg. Top. Life Sci.* ETL520180147 (2019). doi:10.1042/ETLS20180147
258. Filipovich, A. H. *et al.* Impact of donor type on outcome of bone marrow transplantation for Wiskott-Aldrich syndrome: Collaborative study of the International Bone Marrow Transplant Registry and the National Marrow Donor Program. *Blood* **97**, 1598–1603 (2001).
259. Pai, S. Y. *et al.* Stem cell transplantation for the Wiskott-Aldrich syndrome: A single-center experience confirms efficacy of matched unrelated donor transplantation. *Bone Marrow Transplant.* **38**, 671–679 (2006).
260. Hacein-Bey Abina, S. *et al.* Outcome following Gene Therapy in Patients with Severe Wiskott-Aldrich Syndrome HHS Public Access Author manuscript. *Jama* **313**, 1550–1563 (2015).
261. Schnütgen, F. *et al.* High Efficiency Gene Correction in Hematopoietic Cells By Template-Free Crispr/Cas9 Genome Editing. *Blood* **128**, 3507 LP – 3507 (2016).
262. Guo, T. *et al.* Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. *Genome Biol.* **19**, 1–20 (2018).
263. Frecha, C. *et al.* Improved lentiviral vectors for Wiskott-Aldrich syndrome gene therapy mimic endogenous expression profiles throughout haematopoiesis. *Gene Ther.* (2008). doi:10.1038/gt.2008.20
264. Toscano, M. G. *et al.* Was cDNA Sequences Modulate Transgene Expression of Was Promoter-Driven Lentiviral Vectors. *Hum. Gene Ther.* **20**, 1279–1290 (2009).
265. Zhan, J., Johnson, I. M., Wielgosz, M. & Nienhuis, A. W. The identification of hematopoietic-specific regulatory elements for WASp gene expression. *Mol. Ther. - Methods Clin. Dev.* **3**, 16077 (2016).
266. Joglekar, A. V. *et al.* Integrase-defective lentiviral vectors as a delivery platform for targeted modification of adenosine deaminase locus. *Mol. Ther.* **21**, 1705–1717 (2013).
267. Genovese, P. *et al.* Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* (2014). doi:10.1038/nature13420
268. Bak, R. O., Dever, D. P. & Porteus, M. H. CRISPR/Cas9 genome editing in human hematopoietic stem cells. *Nat. Protoc.* (2018). doi:10.1038/nprot.2017.143

ANEXO I: CONSTRUCCIONES

Construcción	Tipo	Elementos que incorpora
Lg ₁ Cas9	Plásmido lentiviral	Sistema CRISPR/Cas9 gl ₁ - Cassette de expresión SFFV-Cas9 - Cassette de expresión U6-gl ₁
Lg ₉ Cas9	Plásmido lentiviral	Sistema CRISPR/Cas9 gl ₉ - Cassette de expresión SFFV-Cas9 - Cassette de expresión U6-gl ₉
Lg ₉ Cas9_IS2	Plásmido lentiviral	Sistema CRISPR/Cas9 gl ₉ - Cassette de expresión SFFV-Cas9 - Cassette de expresión U6-gl ₉ - Elemento IS2 en el 3'LTR
SE	Plásmido lentiviral	- Cassette de expresión SFFV-eGFP (SE)
SEWP	Plásmido lentiviral	- Cassette de expresión SE - Regulador postranscripcional WPRE (WP)
SE-IS2	Plásmido lentiviral	- Cassette de expresión SE - Elemento IS2 en el 3'LTR
SEWP-IS2	Plásmido lentiviral	- Cassette de expresión SE - Regulador postranscripcional WPRE (WP) - Elemento IS2 en el 3'LTR
SE-1.2Kb	Plásmido lentiviral	- Cassette de expresión SE - Fragmento irrelevante de 1.2Kb en el 3'LTR
AWE	Plásmido lentiviral	- Cassette de expresión de eGFP a través del promotor de <i>WAS</i> (AWE)
AWE-IS2	Plásmido lentiviral	- Cassette de expresión AWE - Elemento IS2 en el 3'LTR

SE- δ GF-I ₁ -P	Plásmido lentiviral	- Cassette SE con eGFP mutado mediante la eliminación del primer ATG y la introducción de un fragmento de 387 pb del primer intrón de <i>WAS</i>
SEED	Plásmido lentiviral	- Cassette de expresión SE - Cassette de expresión EF1 α -dsRed (ED)
SECdR	Plásmido lentiviral	- Cassette de expresión SE - Cassette de expresión CMV-dsRed (CdR), en antisentido respecto a SE
SECdR_IS2	Plásmido lentiviral	- Cassette de expresión SE - Cassette de expresión CMV-dsRed (CdR), en antisentido respecto a SE - Elemento IS2 en el 3'LTR
SECdR_WI ₁ R	Plásmido lentiviral	- Cassette de expresión SE - Cassette de expresión CMV-dsRed (CdR), en antisentido respecto a SE - Fragmento de 387pb de <i>WAS</i> , en antisentido respecto a SE
SEWAS84	Plásmido lentiviral	- Cassette de expresión SE precedido por 84pb del primer intrón de <i>WAS</i>
SdR*eGFP	pUC57	- Cassette de expresión SFFV- dsRed (SdR) - Secuencia codificante para eGFP, incluyendo diferentes mutaciones que impiden su expresión.
pUC57 WAS-SE-g9	pUC57	-Diana para el Sistema CRISPR/Cas9 g ₉ - 760pb de brazo de homología 5' rente al sitio de rotura en el locus <i>WAS</i> - Cassette de expresión SE -760pb de brazo de homología en 3' frente al sitio de rotura en el locus <i>WAS</i>

ANEXO II: PUBLICACIONES

The IS2 Element Improves Transcription Efficiency of Integration-Deficient Lentiviral Vector Episomes

Sabina Sánchez-Hernández,^{1,8} Alejandra Gutierrez-Guerrero,^{1,8} Rocío Martín-Guerra,¹ Marina Cortijo-Gutierrez,¹ María Tristán-Manzano,¹ Sandra Rodríguez-Perales,² Laura Sanchez,¹ Jose Luis Garcia-Perez,¹ Jesus Chato-Astrain,⁴ Ricardo Fernandez-Valades,⁵ Ana Belén Carrillo-Galvez,³ Per Anderson,^{2,3} Rosa Montes,³ Pedro J. Real,^{3,6} Francisco Martín,^{1,2,9} and Karim Benabdellah^{1,2,9}

¹Genomic Medicine Department, GENYO, Centre for Genomics and Oncological Research, Pfizer-University of Granada-Andalusian Regional Government, PTS Granada, 18016 Granada, Spain; ²LentiStem Biotech, GENYO, Avda. de la Ilustración 114, 18016 PTS Granada, Spain; ³Oncology Department, GENYO, Centre for Genomics and Oncology, Pfizer-University of Granada-Andalusian Regional Government, PTS Granada, 18016 Granada, Spain; ⁴Department of Histology, Tissue Engineering Group, University of Granada, Granada, Spain; ⁵Pediatric Surgery Department, University Hospital "Virgen de las Nieves," Avda. Fuerzas Armadas 2, 18014 Granada, Spain; ⁶Department of Biochemistry and Molecular Biology I, University of Granada, Granada, Spain; ⁷Molecular Cytogenetics and Genome Editing Unit, Human Cancer Genetics Department, CNIO, Melchor Fernandez Almagro 3, 28029 Madrid, Spain

Integration-defective lentiviral vectors (IDLVs) have become an important alternative tool for gene therapy applications and basic research. Unfortunately, IDLVs show lower transgene expression as compared to their integrating counterparts. In this study, we aimed to improve the expression levels of IDLVs by inserting the IS2 element, which harbors SARs and HS4 sequences, into their LTRs (SE-IS2-IDLVs). Contrary to our expectations, the presence of the IS2 element did not abrogate epigenetic silencing by histone deacetylases. In addition, the IS2 element reduced episome levels in IDLV-transduced cells. Interestingly, despite these negative effects, SE-IS2-IDLVs outperformed SE-IDLVs in terms of percentage and expression levels of the transgene in several cell lines, including neurons, neuronal progenitor cells, and induced pluripotent stem cells. We estimated that the IS2 element enhances the transcriptional activity of IDLV LTR circles 6- to 7-fold. The final effect the IS2 element in IDLVs will greatly depend on the target cell and the balance between the negative versus the positive effects of the IS2 element in each cell type. The better performance of SE-IS2-IDLVs was not due to improved stability or differences in the proportions of 1-LTR versus 2-LTR circles but probably to a re-positioning of IS2-episodes into transcriptionally active regions.

INTRODUCTION

Lentiviral vectors (LVs) have proven to be highly successful in several gene therapy protocols over the last 20 years.¹ Their success is partly explained by their ability to transduce dividing and non-dividing cells, including hematopoietic stem cells (HSCs),²⁻⁴ neurons,⁵⁻⁷ and T cells.⁸⁻¹⁰ However, concerns have been raised regarding the potential risk of inherent insertional mutagenesis caused by integrative LVs.¹¹ The use of integration-defective LVs (IDLVs) is a logical op-

tion for minimizing insertional mutagenesis risk when target cells are quiescent. IDLVs are also an interesting alternative for transient expression in dividing cells.¹²⁻²⁰ As with their integrative counterpart, the tropism of IDLV particles can be altered and adapted to target cells pseudotyped with different envelope proteins.^{21,22}

Current research in the field of retroviral vectors shows that LVs are often more effective than their IDLV counterparts in terms of gene expression^{23,24} mainly due to the tendency of IDLVs to undergo epigenetic silencing as a result of nuclear chromatinization.^{25,26} In 2013, Pelascini et al.²⁷ showed that histone deacetylase (HDAC) activity is the principal cellular determinant underlying weak IDLV transcriptional activity. Different strategies have been used to enhance the transgene expression of episomal molecules in both viral- and non-viral-based systems. In non-viral episomal gene delivery systems, genomic elements based on scaffold or matrix attachment regions (SARs or S/MARs) are widely used to enhance transcription levels and to maintain long-term expression rates.²⁸⁻⁴¹ This is mainly due to the capacity of SARs elements to bind transcription factors such

Received 12 January 2018; accepted 14 August 2018;
<https://doi.org/10.1016/j.omtn.2018.08.007>.

⁸These authors contributed equally to this work.

⁹Senior author

Correspondence: Francisco Martín, Genomic Medicine Department, GENYO, Centre for Genomics and Oncological Research, Pfizer-University of Granada-Andalusian Regional Government, PTS Granada, Health Sciences Technology Park, Avda. de la Ilustración 114, 18016 Granada, Spain

E-mail: francisco.martin@genyo.es

Correspondence: Karim Benabdellah, Genomic Medicine Department, GENYO, Centre for Genomics and Oncological Research, Pfizer-University of Granada-Andalusian Regional Government, PTS Granada, Health Sciences Technology Park, Avda. de la Ilustración 114, 18016 Granada, Spain

E-mail: karim.benabdel@genyo.es



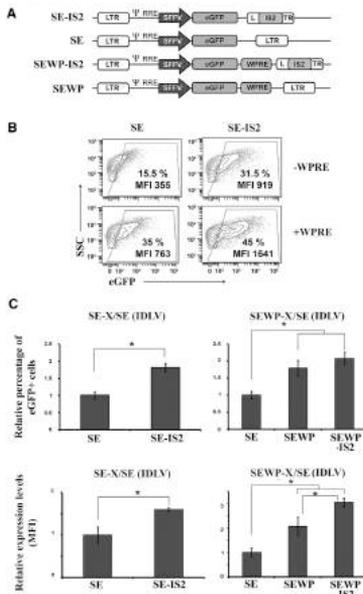


Figure 1. Inclusion of IS2 Element into IDLVs Enhances eGFP Expression Levels in 293T Cells

(A) Schematic representation of SE-IS2, SE, SEWP-IS2, and SEWP. eGFP, enhanced green fluorescence protein; SFV, spleen focus forming virus promoter; WPRE, woodchuck hepatitis virus posttranscriptional regulatory element. (B) Representative plots showing eGFP expression profiles of 293T cells transduced with the different IDLVs. An MOI of 0.3 was used to maintain the percentage of eGFP⁺ cells below 50%. The eGFP⁺ population gates were set to 0.2%–0.7% of eGFP⁺ cells in the untransduced population and subtracted from the % obtained under the different vectors and conditions for the analysis. The percentages (%) and expression levels (MFI) of the eGFP⁺ population are shown in each plot. (C) Graphs showing relative % of GFP⁺ cells (top graphs) and relative expression levels (MFI, bottom graphs) in 293T cells of SE-IS2-IDLVs and SE-IDLVs in the absence (left graphs) or presence (right graphs) of the WPRE element. Values represent means ± SEM of at least four separate experiments (*p < 0.05).

as special AT-rich binding protein 1 (SATB1), nuclear matrix protein 4 (Nmp4) and CCCTC-binding factor (CTCF), in addition to their capacity to promote nucleoprotein structural aggregation, histone acetyltransferase recruitment, and ATP-dependent chromatin re-

modeling complexes.⁴² As with the SAR elements, the inclusion of other *cis*-acting elements based on the 5' cHS4 chicken hypersensitive site 4 (HS4) chicken β -globin insulator (cHS4) also increases non-viral episomal efficiency³³ due, in part, to the interaction of the cHS4 element with the matrix via CTCF nuclear proteins.⁴³ This nuclear factor avoids heterochromatin spreading in episomal DNA when bound to the cHS4 sequence.^{44–47}

Several research groups have achieved varying levels of success in their attempt to improve IDLV expression by inserting different fragments of the β -interferon SARs element into IDLVs.^{48–50} However, the use of SARs elements to improve IDLV transcription efficiency has only been studied in relation to immunoglobulin κ (Igc) SARs sequences. Grandchamp et al.⁵¹ concluded that these elements had, in general, no effect on transduction efficiency and observed an improvement only in differentiated primary neural progenitor cells.

We have previously reported that the inclusion of the IS2 element (which combines a synthetic SAR [SAR2] and a 650-pb fragment of the chicken β -globin HS4 insulator) in LVs reduced biological viral titers but improved transgene expression and prevents epigenetic silencing in human embryonic stem cells (hESCs) and hematopoietic pluripotent cells (HSCs). However, these effects were cell type dependent, since no improvement in transgene expression could be observed on K562 cells and other immortalized cell lines.⁵² In the present study, we tested whether the IS2 element could improve IDLV gene expression in different cell types. Although the presence of the IS2 element did not abrogate epigenetic silencing, it did improve IDLV efficiency in several cell lines. Surprisingly, in spite of the improved expression levels, the inclusion of the IS2 element into IDLVs (SE-IS2-IDLVs) reduced 3–5 times the amount of episomal vector in transduced cells relative to those transduced with unmodified SE-IDLVs. We have estimated that the IS2 element enhances the transcriptional activity of SE-IS2-episomes 6- to 7-fold. The final effect of the IS2 element in IDLVs will greatly depend on the target cell and the balance between the negative versus the positive effects of the IS2 element in each cell type. We have also shown that the IS2 element does not improve the stability of IDLV episomes and that, although there is a change in the ratio of 1-LTR/2-LTR circles, this is not the mechanism involved in the increased transcriptional activity of the SE-IS2-IDLV episomes. Finally, a fluorescence *in situ* hybridization (FISH) analysis suggested that the improved behavior SE-IS2-IDLV episomes is probably due to a distinct nuclear re-positioning into transcriptionally active regions, as suggested by the aggregation of SE-IS2-IDLV episomes into DAPI-low regions.

RESULTS

The Inclusion of the IS2 Insulator in the Long Terminal Repeat of IDLVs Improves Their Expression Levels in 293T Cells in an HDAC-Independent Manner

We generated IDLV particles from an SE lentiviral backbone containing or not the IS2 element^{52–55} with and without WPRE (woodchuck hepatitis virus posttranscriptional regulatory element) (Figure 1A). We first analyzed the efficiency of different IDLVs in 293T cells.

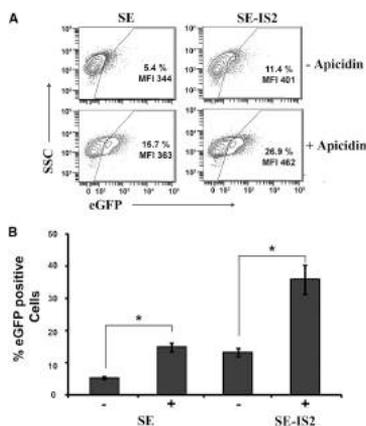


Figure 2. Apicidin Enhances Gene Expression of IDLV Transduced Cells Independently of the Presence of the IS2 Element

(A) Representative plots showing eGFP expression profiles of 293T cells transduced with the SE or SE-IS2 at MOI = 0.2, in the absence or presence of 0.4 μ M apicidin. (B) Graphs showing % of GFP⁺ cells transduced with the SE-IS2-IDLVs and SE-IDLVs in the absence (-) or presence (+) of 0.4 μ M apicidin. Values represent means \pm SEM of at least four separate experiments ($p < 0.05$).

These cells were transduced with an equal MOI, estimated based on the Applied Biological Materials (ABM) Lentiviral qPCR Titer Kit (see [Materials and Methods](#)), and 3 days later, we analyzed the percentage of eGFP⁺ cells and the transgene expression levels (measured as mean fluorescence intensity [MFI] of the eGFP⁺ population). We found that the incorporation of the IS2 element into the IDLVs significantly increased the expression levels of eGFP in the absence and presence of the WPRE element (Figure 1B, MFI; Figure 1C, bottom graphs). We also found an increase in the percentage of GFP⁺ cells, which reached significance only in the absence of the WPRE element (Figure 1B, %; Figure 1C, upper graphs). We further corroborated that the effect of the IS2 element on IDLVs was maintained at higher MOIs (Figure S1).

Prevention of histone deacetylation, the main factor underlying weak IDLV transcriptional activity, could explain the higher SE-IS2-IDLVs expression levels.²⁷ In order to study this possibility, we analyzed SE-IDLV and SE-IS2-IDLV GFP expression levels in the presence and absence of apicidin, an HDAC inhibitor (HDACi). As can be observed in Figure 2, the addition of apicidin enhanced the eGFP expression to a similar degree in cells transduced with SE-IDLVs and in those transduced with SE-IS2-IDLVs

(2.90-fold and 2.35-fold, respectively). These findings suggest that IS2-mediated enhancement is caused by an HDAC-independent mechanism.

The Insertion of the IS2 Element into the IDLV Backbone Does Not Affect RNA Packaging into Vector Particles but Reduces the Amount of IDLV Episomes in the Target Cells

It has been described previously that insertions of large fragments into the 3' long terminal repeat (LTR) of LVs reduce their efficacy. These insertions do not affect viral particle production but reduce reverse transcription efficacy in target cells.^{53,55,56} We therefore analyzed the effect of IS2 (1.2 kb long) on LV and IDLV transgene expression, vector production, and reverse transcription products in 293T target cells. The vector production efficacy was measured from the vector supernatants using ABM's Lentiviral qPCR Titer Kit. This kit calculates the transduction units (TU) per mL based on an equation that convert genome copies per mL (GC/mL) values to TU/mL (see [Materials and Methods](#) for details). The value obtained using this formula is generally very close to the amount of effective particles per mL,⁵⁷ although this can vary depending on the vector backbone. The relative amounts of reverse transcription products of the different vectors were quantified by qPCR using the U3Fw/PBSRev primers pair (see Figure S2). As expected, we did not observe any effect of the IS2 on vector production of LVs or IDLVs (data not shown). Also in agreement with previous observations,^{52–55} the insertion of IS2 into the LTR had a negative effect on the expression levels of integrative LVs (Figure 3A, left plots) that correlated with a reduction of the amount of reverse transcribed products in target cells (Figure 3B, left bars). Interestingly, although the insertion IS2 also caused a similar reduction of IDLV reverse transcription products (Figure 3B, right bars), we observed a significant improvement in both the percentage and the expression levels (Figure 3A, right plots).

IDLV Episomes Harboring the IS2 Element Express Higher mRNA Levels and Have a Distinct Nuclear Localization

The IS2 is a chimeric DNA element containing a synthetic SAR (SAR2) and a 650-pb fragment of the chicken β -globin HS4 insulator.⁵² As mentioned in the introduction, these elements could enhance transcription or increase episomal stability^{26–29,41} that can also lead to an improved transgene expression. In order to study the mechanisms involved in the IS2 effect, we first analyzed the increment in transcriptional efficacy of SE-IS2 episomes in comparison with SE episomes. To do that, we measured eGFP mRNA expression levels of 293T cells transduced with equivalent MOIs of SE-IDLVs and SE-IS2-IDLVs at 72 hr post-transduction and normalized to the amount of vector genomes (Figure 4A). This analysis showed that the SE-IS2 episomes express 6–7 times more mRNA than SE episomes.

We next analyzed whether the enhanced transcriptional activity of SE-IS2 episomes was due to an effect on their longtime stability. We therefore studied the eGFP expression levels (Figure 4B) and the relative amount of SE-IS2 episomes (Figure 4C) related to SE

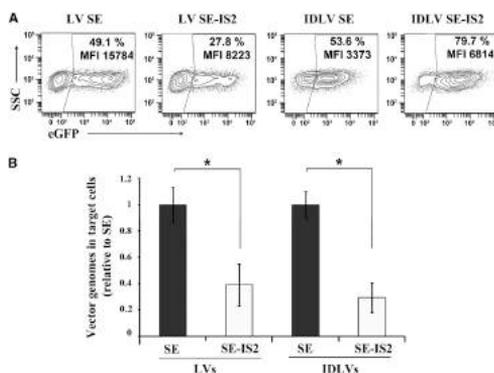


Figure 3. The Insertion of the IS2 Element into the LV Backbone Reduces Their Efficiency to Generate Vector Genomes in Target Cells

(A) Representative plots showing eGFP expression profiles of 293T cells transduced with integrative LVs (left plots) and IDLVs (right plots) with (SE-IS2) or without (SE) the IS2 element. All the experiments were carried out using 0.7 viral particle/cells. (B) Graph showing the amounts of reverse-transcribed products (vector genomes) in 293T cells transduced with LVs (left bars) and IDLVs (right bars) with (SE-IS2) or without (SE) the IS2 element at 72 hr after transduction. Values represent means \pm SEM of at least four separate experiments ($p < 0.05$).

episodes at different time points post-transduction (24 hr to 7 days). Our data showed that the IS2 element does not influence the stability of the IDLV expression (Figure 4B) or the stability of IDLV episodes (Figure 4C).

We finally analyzed whether the presence of IS2 in the IDLV episodes could influence their nuclear localization to transcriptionally active sites since SAR elements bind several factors that promote the aggregation of nucleoproteins, HDAC recruitment, and chromatin remodeling complexes that enhance transcription.^{42,44–47} We performed a FISH analysis on 293T cells transduced with SE-IDLVs and SE-IS2-IDLVs at an MOI of 10 using SE plasmid as probe. As we can observe in Figures 4D and S3, the SE-IDLV episodes are uniformly distributed inside the nuclei, while the SE-IS2-IDLV episodes follow a more aggregated pattern. An analysis of the DAPI/probe co-localization indicated that the episodes harboring IS2 localized preferentially into regions with lower DAPI signal compared to the SE-IDLV episodes (Figure 4E). These data indicated that the better behavior of SE-IS2-IDLV episodes is probably due, in part, to a distinct nuclear re-positioning compared to SE-IDLV episodes.

Insertion of the IS2 Element into the IDLV Reduces the Formation of Lower Expressing 2-LTR Circles

Large LTR inserts such as the IS2 element can also increase homologous LTR recombination favoring 1-LTR circle formation that are superior to 2-LTR circles in terms of transgenic expression,^{36,38} and this could be another potential mechanism behind the improved transcriptional activity of SE-IS2-IDLVs. In order to study this possibility, we generated a control IDLV (SE-1.2kb) harboring an equivalent insertion (1.2 kb) of irrelevant DNA at the same LTR location

(Figure 5A, top) and compared its effect on IDLVs behavior. As observed previously with the IS2 element, the insertion of the 1.2-kb fragment did not affect vector titer, estimated based on ABM's Lentiviral qPCR Titer Kit (see Materials and Methods) (Figure 5B) and reduced 4–5 times the amount of reverse-transcribed products in 293T target cells (Figure 5C). Additionally, as expected by their increased LTR size, the amount of 2-LTR circles (calculated by qPCR using the q2-LTRfw/q2-LTRrev primers [see Figure S2]) related to total episodes (calculated by qPCR using the Δ U3fw/PBSrev primers [see Figure S2]) was reduced similarly on SE-IS2 and SE-1.2kb IDLVs (Figure 5E), suggesting a similar increase on 1-LTR circles in both IDLVs. However, contrary to the SE-IS2-IDLVs, the transgene expression of SE-1.2kb IDLVs was 2.5 lower than SE-IDLVs (Figure 5A, plots, and Figure 5D), which correlates with the observed reduction on IDLVs episodes (Figure 5C). These data indicate that the decrease of the 2-LTR circles is not the main mechanism involved on the increased transcriptional activity of the SE-IS2-IDLV episodes, since the 293T cells transduced with the SE-1.2kb IDLVs showed a similar 2-LTR decrease but no effect on transcription efficacy.

The Final Effect of the IS2 Element Depends on the Target Cell and Vector Backbone

We next studied the potential applications of IS2-IDLVs in different target cells of interest with regard to gene therapy and/or basic research, such as neural progenitor cells (NPCs), neuronal cells (NCs), induced pluripotent stem cells (iPSCs), mesenchymal stromal cells (MSCs), human oral mucosa (HOM), human skin fibroblasts (HSF), and T cells. The different cell types were transduced with equal amounts of SE-IDLVs/SE-IS2-IDLV (NPCs, MOI = 3; iPSCs, MOI = 10; HSFs and HOMs, MOI = 10; T cells, MOI = 5; hMSCs, MOI = 5) and, 3 days later, we analyzed transgene expression levels in terms of MFI and percentage of eGFP⁺ cells. The incorporation of the IS2 element resulted in a marked increase in GFP expression levels in NPCs regardless of the presence or absence of the WPRE element (MFI SE = 740 versus SE-IS2 = 2,048, and SEWP = 4,410 versus SEWP-IS2 = 10,073) (Figures 6A and 6C). However, in differentiated

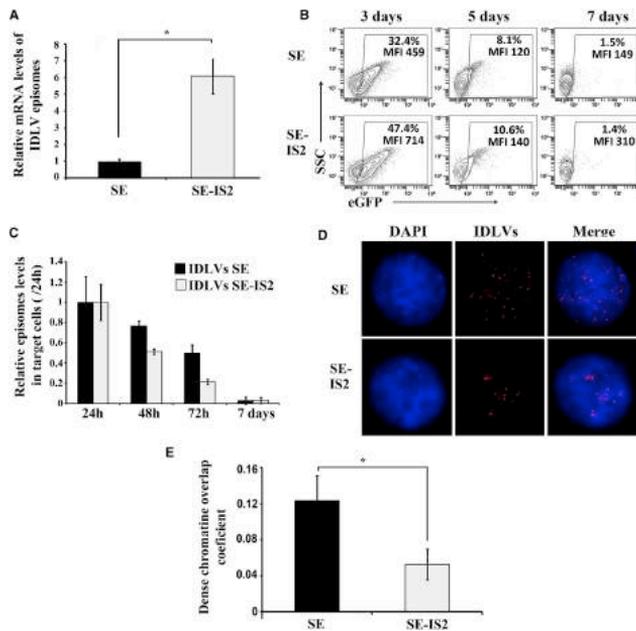


Figure 4. The IS2 Element Does Not Affect the Stability of IDLV Episomes but Enhances Their Transcriptional Efficacy and Relocates Them into DAPI-Low Nuclear Domains

(A) Graph showing the relative amount of eGFP mRNA expression level in 293T cells transduced with IDLV-SE and IDLV-SE-IS2 normalized to the amount of vector genomes in each target cell. (B) Representative plots showing eGFP expression profiles of 293T cells transduced with SE (top) and SE-IS2 (bottom) IDLVs at MOI = 0.5 and analyzed at days 3, 5, and 7 post-transduction. The percentages (%) and expression levels (MFI) of the eGFP⁺ population are shown in each plot. (C) Relative amount of SE-IS2-IDLVs episomes related to SE-IDLVs at different time points post-transduction. (D) Representative confocal images showing nuclear distribution of SE- and SE-IS2-IDLV episomes. The cells were transduced at MOI = 10 in order to obtain several episomes inside the cells. The cells were subsequently fixed, methanol permeabilized, and incubated with Alexa Fluor 555 IDLV-labeled probes. (E) Graph showing the chromatin/IDLV colocalization in 293T cells transduced with SE- and SE-IS2-IDLVs. Values represent the Mander's overlap coefficient (MOC; see [Materials and Methods](#) for details) between DAPI (DNA staining blue) and Alexa Fluor 555 (IDLV staining, in red). Values represent means \pm SEM of at least four separate experiments ($p < 0.05$).

neurons (Figures 6B and 6D; Figure S4), the inclusion of IS2 enhanced expression levels only in WPRE-negative IDLVs (MFI SE = 797 versus SE-IS2 = 2,178), a positive effect that was masked by the inclusion of the WPRE sequence (Figure 6B, bottom; Figure 6D, right bars). As previously shown in undifferentiated neurons, the incorporation of IS2 into the IDLV backbone significantly improved performance in iPSCs either in the absence (Figure 7A, top plots; MFI SE-

IS2 = 894 versus MFI SE = 613) or in the presence of WPRE (Figure 7A, bottom plots; MFI SEWP-IS2 = 1,426 versus MFI SEWP = 1,154; Figure 7B, right graph). Contrary to NPCs and iPSCs, we could not find any significant positive effect of IS2-IDLVs in terms of global effects on eGFP⁺ expression pattern in the different T cell subpopulations (Figure S5), HOM, HSF, or hMSCs (Figure S6 and data not shown).

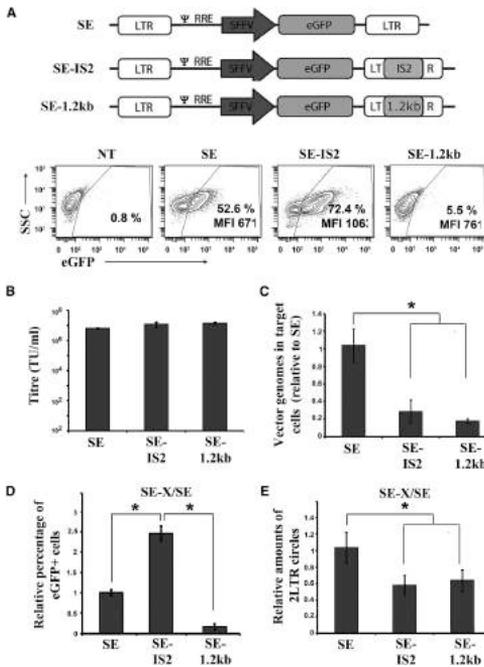


Figure 5. Insertion of the IS2 Element or a Control 1.2-kb Fragment into the IDLVs Reduce the Total Number of Episomes and the Formation of Lower Expressing 2-LTR Circles
 (A) Top, schematic representation of SE, SE-IS2, and SE-1.2kb eGFP; SFFV (spleen focus-forming virus) promoter. Bottom, representative plots showing eGFP expression profiles of 293T cells transduced with the different IDLVs at MOI = 0.7. The percentages (%) and expression levels (MFI) of the eGFP+ population are shown in each plot. (B) Graph showing the titer of the different IDLV supernatants using the ABM's Lentiviral ePCR Titer Kit (see *Materials and Methods* for details). (C) Relative amounts of vector genomes in 293T cells transduced with the different IDLVs at 72 hr post-transduction and normalized to the levels observed in SE IDLVs (see *Materials and Methods* for details). (D) Graph showing the percentage of GFP positive cells in 293T cells transduced with the different IDLVs and normalized to the levels observed in SE-IDLVs. (E) Graph showing the relative amounts of 2-LTR circles relative to total viral genomes in 293T target cells at 72 hr post-transduction and normalized to the levels observed in SE-IDLVs (see *Materials and Methods* for details). The values represent means \pm SEM of at least four separate experiments ($p < 0.05$). The data were reported to an equal amount of cells, estimated by β -albumin house-keeping gene.

DISCUSSION

The success of gene therapy greatly depends on the availability of appropriate gene transfer vectors for the selected strategy. The broad application of gene therapy to different diseases has been possible thanks to the development of a wide range of vectors with different properties.¹ In general, integrative vectors are the vectors of choice when stable gene expression in actively dividing cells is required, whereas non-integrative vectors are preferred for stable expression in non-dividing cells or when transient expression is sufficient or desirable. Integrative LVs

We finally analyzed whether the IS2 element had a similar effect when inserted in a different IDLV backbone expressing the transgene under a physiological promoter. We used an LV backbone previously published by our group (AWE)⁵⁰ that express eGFP through the WAS (Wiskott-Aldrich syndrome) gene promoter only in hematopoietic cells⁶⁰ and generated the IS2-AWE (Figure 8, top). Since the WAS promoter is hematopoietic specific, we tested the behavior of AWE-IDLVs versus AWE-IS2-IDLVs in Jurkat cells (an immortalized line of human T lymphocytes). As can be observed in Figure 8, the results were similar to that obtained with the SE-IDLVs. Indeed, although the cells transduced with AWE-IS2-IDLVs had 3–4 times lower episome numbers on target cells, compared to those transduced with AWE-IDLVs (Figure 8, graph), they expressed similar or slightly higher levels of eGFP (Figure 8, plots).

have several properties that make them an attractive tool for gene delivery, such as the ability to deliver inserts of up to 12 kb, the active translocation to intact nuclei, and the possibility of using different envelopes that enable efficient gene delivery in almost all the cell lines analyzed. These characteristics have led scientists to extend the range of potential applications by developing IDLVs,⁶¹ which maintain several properties of LVs while expressing the transgene without integrating their genome into the host chromosome. IDLVs extend the applicability of LVs, which are safer when stable expression is required in non-dividing cells (neurons, hepatocytes) and can also achieve transient expression in actively dividing cells.^{62,63} However, the expression levels and titers of IDLVs are generally lower than those of LVs,^{64–66} which has limited the applications of this technology.

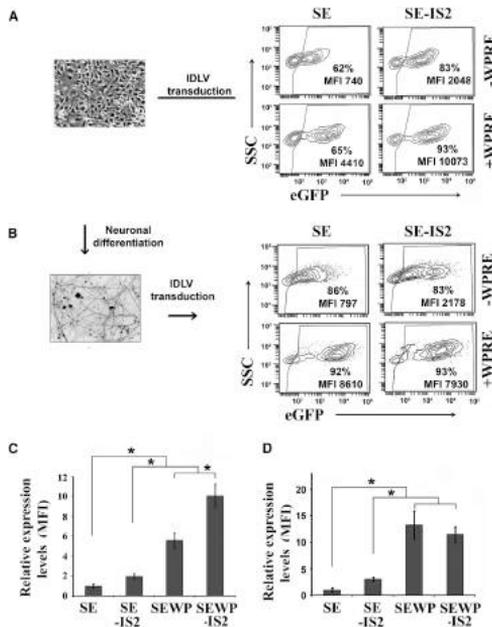


Figure 6. The IS2 Sequence Improves Transgene Expression in Neural Progenitor Cells and Differentiated Neuron-like Cells

(A) Representative image (left) and plots (right) showing eGFP expression profiles of undifferentiated neural stem cells transduced at MOI = 3 with SE, SE-IS2, SEWP, and SEWP-IS2-IDLVs and analyzed 72 hr post-transduction. The percentage (%) of the eGFP⁺ cells and the eGFP expression levels (MFI) are shown in each plot. (B) Representative image (left) and plots (right) showing eGFP expression profiles of differentiated neuron-like cells (positive for β -tubulin, see *Materials and Methods* for details) transduced at MOI = 3 with SE, SE-IS2, SEWP, and SEWP-IS2-IDLVs and analyzed 72 hr post-transduction. The percentage (%) of the eGFP⁺ cells and the eGFP expression levels (MFI) are shown in each plot. (C) Graph showing the relative expression levels of eGFP (MFI) in NPCs transduced with the different IDLVs and normalized to the levels observed in SE-IDLVs. (D) Graph showing the relative expression levels of eGFP (MFI) in differentiated NPCs (neuron-like cells) transduced with the different IDLVs and normalized to the levels observed in SE-IDLVs. Values represent means \pm SEM of at least four separate experiments (* p < 0.05).

HS4 insulator) that avoid silencing and enhance LV expression in hESCs and HSCs.²² We therefore hypothesized that the inclusion of the IS2 element in IDLVs could improve their behavior by avoiding epigenetic transcriptional silencing through HS4 activity and by improving transcription efficiency through SAR activity. Contrary to expectations, the presence of the IS2 element did not abrogate epigenetic silencing by HDACs, although it did improve the transcriptional efficiency of episomal IDLVs.

Up to now, most attempts to improve IDLVs have focused on improving the stability of DNA episomal circles either through transient cell cycle arrest²⁰ or by the inclusion of SAR elements alone^{48–50} or combined with replication origin (<http://www.vivebiotech.com/technology>). However, efficient stable transgene expression in highly dividing cells using IDLVs can be difficult to achieve. As an alternative, we have focused on improving the transient transgene expression levels of IDLVs through insertion of the IS2 element. Although the presence of SAR sequences in this element could also affect expression stability, we did not find this type of effect, which is in line with the observations of Kymäläinen et al.,⁵⁰ who did not observe any differences in episomal establishment in IDLVs containing a SAR sequence. However, these data differ from other studies that show that the insertion of a full 1.2-kb or minimal 155-pb fragment of the β -interferon SAR elements in IDLVs provided a sustained transgene expression.^{48,49} These contradictory findings could be explained by the differences in SAR elements used in the different IDLVs. Verghese et al.⁴⁹ and

The relatively low expression levels of IDLVs have been linked to epigenetic silencing through cellular defense mechanisms that apply heterochromatin marks to episomal viral sequences. This cellular response, which is not restricted to IDLV systems, affects different vector genomes such as herpes simplex viruses and adenoviruses.^{25,26,67,68} The episomal viral DNA is “chromatinized” and acquires nucleosome-like properties immediately after entry into the nucleus.^{26,68} In particular, IDLV genomes have been previously reported to undergo heterochromatinization through histone deacetylation, a process that can be reversed using HDACi such as sodium butyrate²⁷ and valproic acid.²⁴ For basic research purposes and some clinical applications, HDACi could be used to improve IDLV efficiency; however, in most gene therapy settings, the use of HDACi is not desirable due to potential severe side effects, such as the development of malignancies. We have previously described an improved IS2 insulator (which combines a synthetic SAR element [SAR2] and a 650-pb fragment of the chicken β -globin

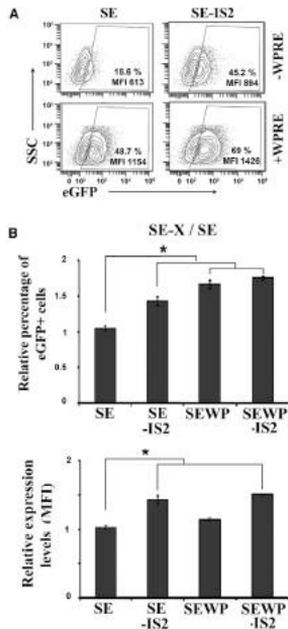


Figure 7. The IS2 Sequence Improves Transgene Expression in iPSCs
 (A) Representative plots showing eGFP expression profiles of iPSCs transduced with the SE, SE-IS2, SEWP, and SEWP-IS2 IDLVs at MOI = 10. The percentage (%) of the eGFP⁺ population and the relative transgene expression levels (MFI) are shown in each plot. (B) Graphs showing the relative percentage of eGFP⁺ cells (top) and the relative transgene expression levels (bottom) in iPSCs transduced with different IDLVs and normalized to the levels obtained with SE-IDLVs. The values represent means \pm SEM of at least four separate experiments ($p < 0.05$).

Xu et al.⁴⁸ used the full β -interferon SARs element and a 155-bp fragment, respectively. On the other hand, Kymäläinen et al.⁵⁰ used a truncated form of 0.7 kb, while the IS2 element contained a synthetic SARs element consisting of 4 SAR recognition signatures (MRS).⁴⁸ It is therefore possible that both the 1.2-kb β -interferon SAR element and the 155-bp fragment are required for episomal maintenance.

We inserted the IS2 element into the lentiviral backbone 3' LTR in order to be duplicated during the reverse transcription process. This

procedure enhances the effect of SAR sequences (the more sequences the better) and/or increases homologous recombination in order to promote the formation of 1-LTR circles, which is reported to be 2- to 4-fold more effective for expression than 2-LTR circles.⁶⁹ To differentiate between these two effects, we constructed a 1.2 kb IDLV lentiviral backbone that has the same insertion in the LTR as IS2-IDLV but contains an irrelevant sequence. The 1-LTR form is the result of homologous recombination between the LTRs, while 2-LTR circles are the result of non-homologous end joining (NHEJ),^{70,71} meaning that longer LTRs are expected to render higher levels of 1-LTRs. Our analysis showed that the inclusion of both the IS2 element and the irrelevant sequence 1.2 kb did not affect vector titer (estimated based on ABM's Lentiviral qPCR Titer Kit; see [Materials and Methods](#)) but reduced the amount of episomes on transduced cells and also the amount of 2-LTR circles related to total episomes. While IS2-IDLVs episomes showed an increased eGFP transgene expression as compared to unmodified IDLVs, the 1.2kb IDLV episomes showed no effect. These data suggest that, as described previously,⁵⁶ the insertion of the IS2 element into the LTRs reduced the reverse transcription process, rendering these IDLVs less efficient in generating IDLV episomes in target cells. Additionally, they also indicated that the enhanced expression of IS2-IDLV episomes cannot be explained by the increment in 1-LTR forms, since this effect is not present in the SE-1.2kb IDLVs.

The improved behavior of IDLVs harboring the IS2 element could be due to an improved transcriptional activity of the IDLV episomes or to improved mRNA stability and/or expression of IS2-containing transcripts. However, the insertion of IS2 into integrative LVs reduced their expression levels in 293T cells, and they also contained the IS2 element in their mRNAs. We can therefore conclude that the better performance of IS2-containing episomes must be due to an effect related to transcription and not to mRNA stability and/or other effects on the mRNA. We therefore focused our attention on trying to understand the potential mechanism involved in this enhanced transcriptional activity of SE-IS2-IDLVs. By normalizing the transcription levels to the relative amount of total vector DNA genomes, we estimated that the transcriptional activity of IS2 episomes is 6- to 7-fold higher than unmodified SE episomes. These positive effects of the IS2 sequence counter-balance the negative effect on episome generation in target cells. Therefore, the final effect the IS2 element in IDLVs will greatly depend on the target cell and the balance between the negative (less efficacy of episome generation) versus the positive (enhanced transcription) effects of the IS2 element in each cell type. Interestingly, the IS2 element still has a similar activity when inserted into a different LV backbone expressing the transgene through the WAS promoter, suggesting that the effect could be independent of the promoter used.

The elements contained in the IS2 (HS4 and SAR sequences) function as DNA anchor points for the chromatin scaffold and organize the chromatin into structural domains that separate different transcriptional units from each other and provide a platform for the assembly of the factors involved in transcription regulation.^{42,72} Several studies

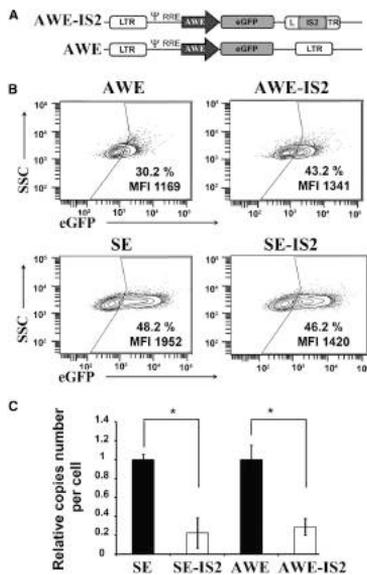


Figure 8. Performance of the IS2-IDLVs Expressing the Transgene through the Human WAS Promoter

(A) Scheme of the IDLVs harboring the human WAS promoter. (B) Representative plots showing eGFP expression profiles of Jurkat cells transduced with AWE and AWE-IS2 (top) and the SE- and SE-IS2 (bottom) IDLVs at MOI = 0.5. The percentage (%) of the eGFP⁺ population and the expression levels (MFI) are shown in each plot. (C) Graphs showing the relative vector DNA genomes in IDLV-transduced cells at 72 hr after transduction. Values represent means \pm SEM of at least four separate experiments ($p < 0.05$).

indicate that these elements are located in proximity to expressed genes at the 5' end or near transcription start sites.⁷² Several pieces of evidence suggest that these elements may poise the DNA for transcription by allowing interaction with ubiquitous tissue-specific transcription factors, such as special SATB-1,^{73,74} NMP4,⁷⁵ and CTCF;⁷⁶ these, in turn, recruit regulatory proteins such topoisomerases and ATP-dependent chromatin remodeling complexes to mediate a more expression-permissive state.⁷⁷ These nuclear domains involved in transcription and replication stain poorly with DAPI due to a lower DNA context.⁷⁸ In this direction, our FISH analysis showed that, while the SE-IDLV episodes were uniformly distributed throughout the

nuclei, the SE-IS2-IDLV episodes followed a more aggregated pattern into DAPI-low regions. These observations indicate that the improved behavior of SE-IS2-IDLV episodes is probably in part due to a preferential nuclear re-positioning into transcriptionally active regions.

In summary, we have designed a new LV backbone that improves the expression levels of IDLVs through a 6- to 7-fold increase in the transcriptional activity of IDLV episodes due to the inclusion of the IS2 element at the LTR. However, as the inclusion of this element also negatively affects the amount of episomal vectors produced in target cells, the final effect on different cell types varies. We did observe an objective increase in transgene expression only in certain cell types, such as 293T cells, NPCs, neurons, and iPSCs. Additional improvements such as the reduction of IS2 element size and/or the design of new backbones harboring this element in different locations could further enhance the effectiveness of IDLVs as a vector system.

MATERIALS AND METHODS

Cell Lines

293T cells (CRL11268; American Type Culture Collection [ATCC], Rockville, MD) were cultured in DMEM (Invitrogen, Edinburgh, Scotland), supplemented with 10% fetal bovine serum (FBS; Invitrogen). The Jurkat T lymphocyte line (TIB-152) (ATCC, Manassas, VA) was grown in RPMI medium supplemented with 10% FBS and Penicillin-Streptomycin Solution (Biowest). Human neural progenitor cells (Gibco human neural stem cells [NSCs], H9 hESC-derived, cat. no. N7800) were maintained in StemPro NSC serum-free medium (SFM) (KnockOut DMEM/F-12, StemPro NSC SFM supplement, basic FGF recombinant protein, and EGF recombinant protein; Gibco, Thermo Fisher Scientific). Plates were previously coated with poly-L-Ornithine 10 μ g/mL (Sigma-Aldrich, St. Louis, MO; <https://www.sigmaaldrich.com>) and laminin 20 μ g/mL (Thermo Fisher Scientific, cat. no. 23017-015). The human iPSC line PBMC1-iPS4F1 were grown as previously described by Montes et al.⁷⁹ All of the human cells used in this study were donated by healthy individuals after informed consent according to the Institutional Guidelines.

Neural Cell Differentiation In Vitro

When the human neural progenitor cells were 90% confluent, the complete growth medium was replaced by the neural differentiation media (NSC SFM Media without basic fibroblast growth factor [FGF] recombinant protein or epidermal growth factor [EGF] recombinant protein) and maintained at 37°C in 5% CO₂ 20%O₂. The medium was changed every 3–4 days for 28 days. The mature neural-like cells were characterized by immunocytochemistry (see [Supplemental Information](#)).

Lentiviral Vector Constructs

The IS2 element, which combine the HS4-650 fragment and the SAR2, were designed and synthesized as previously described.⁵² SE,⁵⁰ SEWP,⁸¹ AWE,⁵⁹ and SEIS2²⁵ plasmids have been previously described. The SEWP-IS2 and AWE-IS2 were generated by the insertion of the IS2 element into the Bbs1 site of the 3' LTR of the SEWP and AWE plasmids, respectively. The SE-1.2kb plasmid was

Table 1. Pairs of Primers Used to Establish the 2-LTR Relative to the Total IDLV DNA Ratio by Real-Time PCR and Estimate the cDNA Amount of eGFP

Primer Pair	Orientation	Sequence	Amplified Forms
q2-LTR	Fw	5'-GCCTCAATAAAGCTTGCCTTG-3'	2-LTR circles
	Rv	5'-TGGGAGTGAATTAGCCCTTCCA-3'	
ΔU3	Fw	5'-GATCTGCTTTTGGCTTGACT-3'	total viral DNA
	Rv	5'-GAGTCTGCGTCGAGAGAGC-3'	
PBS	Fw	5'-GCTGTATCTTGTGGGCTGT-3'	genomic normalization
	Rv	5'-ACTCATGGGAGCTCTGGTTC-3'	
GFP	Fw	5'-GCCGACAACCACTACCT-3'	GFP cDNA
	Rv	5'-CGTCCATGCCGAGAGTGA-3'	
GAPDH	Fw	5'-ATGGGGGAGGTGAAGGTCCG-3'	GAPDH cDNA
	Rv	5'-GGGGTCATTGATGGCAACAATA-3'	

The q2-LTR pair of primer flanks the 2-LTR circle junction. The second pair of primers amplify all forms of reverse transcribed products (bottom). The qhAlb is used as a control for genomic normalization. GFP cDNA primers pair amplifies the eGFP cDNA. GAPDH primers were used for cDNA normalization. FW, forward; Rv, reverse.

generated by inserting a 1.2-kb fragment from SEN (SFV-eGFP-Nluc) plasmid (<https://www.genscript.com>) and our unpublished data) into the SE plasmid.

Viral Production and Titration

LVs and IDLVs were generated by transient transfection of 293T cells using the different vector plasmid together with the packaging plasmid pCMVDR8.91 (http://www.addgene.org/Didier_Trono) for LVs and pCMVDR8.74 (kindly provided by Howe S.J. [UCL Great Ormond Street Institute of Child Health, University College London, London, UK]) for IDLVs, as well as the p-MD-G plasmid encoding the vesicular stomatitis virus (VSV-G) envelope gene (http://www.addgene.org/Didier_Trono). Transfection was performed with the aid of LipoD293 (SigmaGen Laboratories, Ijamsville, MD, USA) according to the manufacturer's instructions, and the supernatants were harvested at 48 hr and 72 hr after transfection. Viral titers were determined by the estimation of TU/mL using ABM's Lentiviral qPCR Titer Kit to convert the amount of viral copies per mL (GC/mL) to TU/mL as indicated by the manufacturer.

Cell Transduction

Concentrated as well as unconcentrated LV and IDLV supernatants were used to transduce different cell lines. Cells were washed with Dulbecco's PBS (1×) (Biowest), counted, and plated on 48-well plates and incubated for 5 hr with different LV and IDLV particles at different MOIs (from 0.3 to 10). Media were changed after 5 hr. NPCs and 293T cells were washed with Dulbecco's PBS (1×) (Biowest), dissociated with TrypLE (Gibco), and plated on 24-well plates in the presence of the fresh viral particles at different TU/cells. The cell line PBMC1-iPS4F1 was incubated for 5 hr on the day of passage with a concentrated virus in the presence of 8 μg/mL Polybrene and 10 μM Y-27632 (Sigma-Aldrich).

Flow Cytometry

At different times (72 hr to 7 days) after transduction, the different cell types were harvested and washed twice with fluorescence-acti-

vated cell sorting (FACS) buffer (PBS containing 2 mM EDTA and 2% FBS), acquired on a FACS Canto II flow cytometer and analyzed using FACS Diva software (BD Biosciences). eGFP expression was detected in the fluorescein isothiocyanate (FITC) channel in a flow cytometer (BD FACScanto).

Profile of Extrachromosomal Forms of IDLV DNA

The 2-LTR:total IDLV DNA ratio in transduced cells was determined by real-time PCR using different primer pairs (Table 1) that enable us to discriminate 2-LTR from total IDLVs DNA. As an internal control, we also used primers for the human albumin locus (hAlb). DNA of transduced cells was extracted 72 hr after transduction using a QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany; <https://www.qiagen.com:443/us>). Real-time PCRs were performed using the QuantiTect SYBRGreen PCR kit (QIAGEN) on a Stratagene MX3005P System (Agilent Technologies, Santa Clara, CA; <https://www.agilent.com>). The PCRs were performed using the following run program: 10' at 95°C for denaturation, 40 cycles of 15 min at 95°C, 60 min at 60°C, and 72°C for 60' followed by the melting curve. PCR data were analyzed according to the comparative C_T method.⁸²

mRNA Analysis by RT-qPCR

Total RNA was obtained using the Trizol reagent (Invitrogen) according to the manufacturer's instructions. RNA samples were reverse transcribed using the Superscript first-strand system (Invitrogen), and qPCRs were performed using the QuantiTect SYBRGreen PCR kit (QIAGEN) on a Stratagene MX3005P system (Agilent Technologies, Santa Clara, CA; <https://www.agilent.com>). The primers used are listed in Table 1.

FISH

FISH based on the use of an orange 8 (Alexa Fluor 555) fluorescent probe was employed to localize IDLVs in cells as previously described.⁸³ In brief, to generate the IDLV-FISH probe, DNA from the vector plasmid was directly labeled by Nick Translation according to the manufacturer's specifications (Invitrogen, Edinburgh,

Scotland). Fixed cells with Carnoy's fixative were hybridized overnight at 37°C with the SE probe. After post-hybridization washes, the cells were counterstained with DAPI in anti-fade solution (Molecular Probes). Images were acquired on a Zeiss LSM 710 confocal microscope (Carl Zeiss, Jena, Germany; <https://www.zeiss.com>). The extent of colocalization was performed using Mander's overlap coefficient (MOC)³⁴ to quantify the degree.

Statistical Analysis

All data are represented as means ± SEM. Statistical analysis was performed using GraphPad Prism software (GraphPad Software, La Jolla, CA; <https://www.graphpad.com>) by applying the unpaired two-tailed t test. Statistical significance was defined as a p value < 0.05.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Materials and Methods and six figures and can be found with this article online at <https://doi.org/10.1016/j.omtn.2018.08.007>.

AUTHOR CONTRIBUTIONS

K.B. and F.M. wrote the manuscript and designed the experiments. A.G.-G., S.R.-P., and S.S.-H. performed the experiments on 293T cells. A.B.C.-G. and P.A. performed the experiments on hMSC cells. J.C.-A. and R.F.-V. performed the experiments skin fibroblast cells. M.T.-M., S.S.-H., and R.M. performed the experiments on T cells. K.B., S.S.-H., L.S., and J.L.G.-P. performed the experiments on NPCs. S.S.-H., R.M., and P.J.R. performed the experiments on iPSCs. K.B., F.M., A.G.-G., and S.S.-H. analyzed the data. All the authors reviewed the manuscript.

CONFLICTS OF INTEREST

The authors have applied for the patent "IS2 element to improved integration-defective lentiviral vectors. Application number: EP18382613.0" Applicant: Fundación Progreso y Salud. This patent application does not alter the authors' adherence to the policies of *Molecular Therapy* — *Nucleic Acids* on sharing data and materials. Therefore, all material presented in this manuscript will be freely available, though not for commercial purposes, to the scientific community.

ACKNOWLEDGMENTS

This study is financed by the ISCIII Health Research Fund (Spain) and the European Regional Development Fund (FEDER) through research grants PI12/01097 and PI15/02015 (F.M.) and CD09/0020, PI15/00794, and CPII15/00032 (P.A.); the ISCIII Cellular Therapy Network (TerCel, RD12/0019/0006) (F.M.); a Juan de la Cierva fellowship (ICL_2012_12666; to R.M.); the CICE and CS of the Junta de Andalucía FEDER/European Cohesion Fund (FSE) for Andalucía 2007-2013 through research grants P09-CTS-04532, PI-57069, PI-0001/2009, and PAIDI-Bio-326 (F.M.) and PI-0160/2012 and PI-0014-2016 (K.B.); the Ministerio de Economía, Industria y competitividad through research grant SAF2017-89745-R (J.L.G.-P.); a RYC contract (RYC-2015-18382; to P/J.R.); an FPU fellowship (FPU16/05467; to M.T.-M.); a PEJ contract (PEJ-2014-A-46314; to

A.B.C.-G. and PEJ-2014-A-17105 S.S.-H.); the European Research Council (ERC-Consolidator ERC-STG-2012-233764); and a private donation from Ms. Francisca Serrano (Trading y Bolsa para Torpes, Granada, Spain). We also wish to thank Michael O'Shea for proof-reading the article. J.L.G.-P's lab is supported by MINECO-FEDER (SAF2017-89745-R).

REFERENCES

- Naldini, L. (2015). Gene therapy returns to centre stage. *Nature* 526, 351–360.
- Ungari, S., Montepeloso, A., Morena, F., Cocchiarella, F., Recchia, A., Martino, S., Gentner, B., Naldini, L., and Biffi, A. (2015). Design of a regulated lentiviral vector for hematopoietic stem cell gene therapy of globoid cell leukodystrophy. *Mol. Ther. Methods Clin. Dev.* 2, 15038.
- Biffi, A., Aubourg, P., and Cartier, N. (2011). Gene therapy for leukodystrophies. *Hum. Mol. Genet.* 20 (R1), R42–R53.
- Biffi, A., Montini, E., Lorioli, L., Cesani, M., Fumagalli, F., Plati, T., Baldoli, C., Martino, S., Calabria, A., Canale, S., et al. (2013). Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* 341, 1233158.
- Kato, S., Kobayashi, K., and Kobayashi, K. (2014). Improved transduction efficiency of a lentiviral vector for neuron-specific retrograde gene transfer by optimizing the junction of fusion envelope glycoprotein. *J. Neurosci. Methods* 227, 151–158.
- Allodi, L., Mecollari, V., González-Pérez, F., Eggers, R., Hoyng, S., Verhaagen, J., Navarro, X., and Udina, E. (2014). Schwann cells transduced with a lentiviral vector encoding Fgf-2 promote motor neuron regeneration following sciatic nerve injury. *Glia* 62, 1736–1746.
- Towne, C., and Aebischer, P. (2009). Lentiviral and adeno-associated vector-based therapy for motor neuron disease through RNAi. *Methods Mol. Biol.* 555, 87–108.
- Zhou, Q., Uhlir, K.M., Muth, A., Kimpel, J., Lévy, C., Münch, R.C., Seifried, J., Pfeiffer, A., Thibault, A., Gouilhaly, C., et al. (2015). Exclusive Transduction of Human CD4+ T Cells upon Systemic Delivery of CD4-Targeted Lentiviral Vectors. *J. Immunol.* 195, 2493–2501.
- Amirache, F., Lévy, C., Costa, C., Mangot, P.E., Torbett, B.E., Wang, C.X., Nègre, D., Cosset, F.L., and Verhoeven, E. (2014). Mystery solved: VSV-G-LVs do not allow efficient gene transfer into unstimulated T cells, B cells, and HSCs because they lack the LDL receptor. *Blood* 123, 1422–1424.
- Humbert, J.M., Frecha, C., Amirache Bouafia, F., N'Guyen, T.H., Boni, S., Cosset, F.L., Verhoeven, E., and Halary, F. (2012). Measles virus glycoprotein-pseudotyped lentiviral vectors are highly superior to vesicular stomatitis virus G pseudotypes for genetic modification of monocyte-derived dendritic cells. *J. Virol.* 86, 5192–5203.
- Cesana, D., Ranzani, M., Volpin, M., Bartholomae, C., Duros, C., Artus, A., Merella, S., Benedicenti, F., Sergi Sergi, L., Sanvito, F., et al. (2014). Uncovering and dissecting the genotoxicity of self-inactivating lentiviral vectors in vivo. *Mol. Ther.* 22, 774–785.
- Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provasi, E., Colombo, D.F., Neri, M., Magnani, Z., Cantore, A., Lo Riso, P., et al. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* 8, 861–869.
- Lombardo, A., Genovese, P., Beausejour, C.M., Colleonì, S., Lee, Y.L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D., et al. (2007). Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.* 25, 1298–1306.
- Negri, D.R., Rossi, A., Blasi, M., Michelini, Z., Leone, P., Chiarot, M.V., Baroncelli, S., Perrella, G., Cimarelli, A., Klotman, M.E., and Cara, A. (2012). Simian immunodeficiency virus-Vpx for improving integrase defective lentiviral vector-based vaccines. *Retrovirology* 9, 69.
- Negri, D.R., Michelini, Z., Bona, R., Blasi, M., Filati, P., Leone, P., Rossi, A., Franco, M., and Cara, A. (2011). Integrase-defective lentiviral-vector-based vaccine: a new vector for induction of T cell immunity. *Expert Opin. Biol. Ther.* 11, 739–750.
- Michelini, Z., Negri, D.R., Baroncelli, S., Spada, M., Leone, P., Bona, R., Klotman, M.E., and Cara, A. (2009). Development and use of HIV-based integrase defective lentiviral vector for immunization. *Vaccine* 27, 4622–4629.

17. Hoban, M.D., Cost, G.J., Mendel, M.C., Romero, Z., Kaufman, M.L., Joglekar, A.V., Ho, M., Lumaquin, D., Gray, D., Lill, G.R., et al. (2015). Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood* *125*, 2597–2604.
18. Gutierrez-Guerrero, A., Sánchez-Hernández, S., Galvani, G., Pinedo-Gomez, I., Martín-Guerra, R., Sánchez-Gilbert, A., Aguilar-González, A., Cobo, M., Gregory, P., Holmes, M., et al. (2018). Comparison of Zinc Finger Nucleases Versus CRISPR-Specific Nucleases for Genome Editing of the Wiskott-Aldrich Syndrome Locus. *Hum. Gene Ther.* *29*, 366–380.
19. Ortinski, P.L., O'Donovan, B., Dong, X., and Kantor, B. (2017). Integrase-Deficient Lentiviral Vector as an All-in-One Platform for Highly Efficient CRISPR/Cas9-Mediated Gene Editing. *Mol. Ther. Methods Clin. Dev.* *5*, 153–164.
20. Coluccio, A., Miselli, F., Lombardo, A., Marconi, A., Malagoli Tagliazucchi, G., Gonçalves, M.A., Pinelli, C., Maruggi, G., Del Rio, M., Naldini, L., et al. (2013). Targeted gene addition in human epithelial stem cells by zinc-finger nuclease-mediated homologous recombination. *Mol. Ther.* *21*, 1695–1704.
21. Steffen, L., and Simmons, G. (2016). Pseudotyping Viral Vectors With Emerging Virus Envelope Proteins. *Curr. Gene Ther.* *16*, 47–55.
22. Cronin, J., Zhang, X.Y., and Reiser, J. (2005). Altering the tropism of lentiviral vectors through pseudotyping. *Curr. Gene Ther.* *5*, 387–398.
23. Joglekar, A.V., Hollis, R.P., Kufnec, G., Senadheera, S., Chan, R., and Kohn, D.B. (2013). Integrase-defective lentiviral vectors as a delivery platform for targeted modification of adenosine deaminase locus. *Mol. Ther.* *21*, 1705–1717.
24. Joglekar, A.V., Stein, L., Ho, M., Hoban, M.D., Hollis, R.P., and Kohn, D.B. (2014). Dissecting the mechanism of histone deacetylase inhibitors to enhance the activity of zinc finger nucleases delivered by integrase-defective lentiviral vectors. *Hum. Gene Ther.* *25*, 599–608.
25. Takacs, M., Banati, F., Koroknai, A., Segesdi, J., Salamon, D., Wolf, H., Niller, H.H., and Minarovits, J. (2010). Epigenetic regulation of latent Epstein-Barr virus promoters. *Biochim. Biophys. Acta* *1799*, 228–235.
26. Knappe, D.M., Lieberman, P.M., Jung, J.U., McBride, A.A., Morris, K.V., Ott, M., Margolis, D., Nieto, A., Nevels, M., Parks, R.J., and Kristie, T.M. (2013). Snapshots: chromatin control of viral infection. *Virology* *435*, 141–156.
27. Pelascini, L.P., Maggio, I., Liu, J., Holkers, M., Cathomen, T., and Gonçalves, M.A. (2013). Histone deacetylase inhibition rescues gene knockout levels achieved with integrase-defective lentiviral vectors encoding zinc-finger nucleases. *Hum. Gene Ther. Methods* *24*, 399–411.
28. Mizutani, A., Kikkawa, E., Matsuno, A., Shigenari, A., Okinaga, H., Murakami, M., Ishida, H., Tanaka, M., and Inoko, H. (2013). Modified S/MAR episomal vectors for stably expressing fluorescent protein-tagged transgenes with small cell-to-cell fluctuations. *Anal. Biochem.* *443*, 113–116.
29. Haase, R., Magnusson, T., Su, B., Kopp, F., Wagner, E., Lipps, H., Baiker, A., and Ogris, M. (2013). Generation of a tumor- and tissue-specific episomal non-viral vector system. *BMC Biotechnol.* *13*, 49.
30. Argyros, O., Wong, S.P., Gowes, K., and Harbottle, R.P. (2012). Genetic modification of cancer cells using non-viral, episomal S/MAR vectors for in vivo tumour modelling. *PLoS ONE* *7*, e47920.
31. Wong, S.P., Argyros, O., Coutelle, C., and Harbottle, R.P. (2011). Non-viral S/MAR vectors replicate episomally in vivo when provided with a selective advantage. *Gene Ther.* *18*, 82–87.
32. Argyros, O., Wong, S.P., Niceta, M., Waddington, S.N., Howe, S.J., Coutelle, C., Miller, A.D., and Harbottle, R.P. (2008). Persistent episomal transgene expression in liver following delivery of a scaffold/matrix attachment region containing non-viral vector. *Gene Ther.* *15*, 1593–1605.
33. Hagedorn, C., Antoniou, M.N., and Lipps, H.J. (2013). Genomic cis-acting Sequences Improve Expression and Establishment of a Nonviral Vector. *Mol. Ther. Nucleic Acids* *2*, e118.
34. Hagedorn, C., Wong, S.P., Harbottle, R., and Lipps, H.J. (2011). Scaffold/matrix attachment region-based nonviral episomal vectors. *Hum. Gene Ther.* *22*, 915–923.
35. Sgourou, A., Routledge, S., Spathas, D., Athanassiadou, A., and Antoniou, M.N. (2009). Physiological levels of HBB transgene expression from S/MAR element-based replicating episomal vectors. *J. Biotechnol.* *143*, 85–94.
36. Giannakopoulos, A., Stavrou, E.F., Zarkadis, I., Zoumbos, N., Thrasher, A.J., and Athanassiadou, A. (2009). The functional role of S/MARs in episomal vectors as defined by the stress-induced destabilization profile of the vector sequences. *J. Mol. Biol.* *387*, 1239–1249.
37. Lufino, M.M., Manservigi, R., and Wade-Martins, R. (2007). An S/MAR-based infectious episomal genomic DNA expression vector provides long-term regulated functional complementation of LDLR deficiency. *Nucleic Acids Res.* *35*, e98.
38. Zhao, C.P., Guo, X., Chen, S.J., Li, C.Z., Yang, Y., Zhang, J.H., Chen, S.N., Jia, Y.L., and Wang, T.Y. (2017). Matrix attachment region combinations increase transgene expression in transfected Chinese hamster ovary cells. *Sci. Rep.* *7*, 42805.
39. Wang, T.Y., Yang, R., Qin, C., Wang, L., and Yang, X.J. (2008). Enhanced expression of transgene in CHO cells using matrix attachment region. *Cell Biol. Int.* *32*, 1279–1283.
40. Girod, P.A., Nguyen, D.Q., Calabrese, D., Puttini, S., Grandjean, M., Martinet, D., Regamey, A., Saugy, D., Beckmann, J.S., Bucher, P., and Mermoud, N. (2007). Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. *Nat. Methods* *4*, 747–753.
41. Girod, P.A., Zahn-Zabal, M., and Mermoud, N. (2005). Use of the chicken lysozyme 5' matrix attachment region to generate high producer CHO cell lines. *Biotechnol. Bioeng.* *91*, 1–11.
42. Wang, T.Y., Han, Z.M., Chai, Y.R., and Zhang, J.H. (2010). A mini review of MAR-binding proteins. *Mol. Biol. Rep.* *37*, 3553–3560.
43. Yusufzai, T.M., and Felsenfeld, G. (2004). The 5' HS4 chicken beta-globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc. Natl. Acad. Sci. USA* *101*, 8620–8624.
44. Zielke, K., Full, F., Teufert, N., Schmidt, M., Müller-Fleckenstein, I., Alberter, B., and Enser, A. (2012). The insulator protein CTCF binding sites in the *orf73/LANA* promoter region of herpesvirus saimiri are involved in conferring episomal stability in latently infected human T cells. *J. Virol.* *86*, 1862–1873.
45. Holohan, E.E., Kwong, C., Adryan, B., Bartkuhn, M., Herold, M., Renkawitz, R., Russell, S., and White, R. (2007). CTCF genomic binding sites in *Drosophila* and the organisation of the bithorax complex. *PLoS Genet.* *3*, e112.
46. Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heini, N., Bardenhagen, I., Leers, J., White, R.A., Renkawitz-Pohl, R., Saunweber, H., and Renkawitz, R. (2007). The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J.* *26*, 4203–4214.
47. Herold, M., Bartkuhn, M., and Renkawitz, R. (2012). CTCF: insights into insulator function during development. *Development* *139*, 1045–1057.
48. Xu, Z., Chen, F., Zhang, L., Lu, J., Xu, P., Liu, G., Xie, X., Mu, W., Wang, Y., and Liu, D. (2016). Non-integrating lentiviral vectors based on the minimal S/MAR sequence retain transgene expression in dividing cells. *Sci. China Life Sci.* *59*, 1024–1033.
49. Verghese, S.C., Goloviznina, N.A., Skinner, A.M., Lipps, H.J., and Kurre, P. (2014). S/MAR sequence confers long-term mitotic stability on non-integrating lentiviral vector episomes without selection. *Nucleic Acids Res.* *42*, e53.
50. Kymäläinen, H., Appelt, J.U., Giordano, F.A., Davies, A.F., Ogilvie, C.M., Ahmed, S.G., Laufs, S., Schmidt, M., Bode, J., Yáñez-Muñoz, R.J., and Dickson, G. (2014). Long-term episomal transgene expression from mitotically stable integration-deficient lentiviral vectors. *Hum. Gene Ther.* *25*, 438–442.
51. Grandchamp, N., Henriot, D., Philippe, S., Amar, L., Ursulet, S., Serguera, C., Mallet, J., and Sarkis, C. (2011). Influence of insulators on transgene expression from integrating and non-integrating lentiviral vectors. *Genet. Vaccines Ther.* *9*, 1.
52. Benabdellah, K., Gutierrez-Guerrero, A., Cobo, M., Muñoz, P., and Martín, F. (2014). A chimeric HS4-SAR insulator (HS2) that prevents silencing and enhances expression of lentiviral vectors in pluripotent stem cells. *PLoS ONE* *9*, e84268.
53. Ramezani, A., Hawley, T.S., and Hawley, R.G. (2003). Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator. *Blood* *101*, 4717–4724.
54. Arumugam, P.L., Scholes, J., Prehman, N., Xia, P., Yee, J.K., and Malik, P. (2007). Improved human beta-globin expression from self-inactivating lentiviral vectors carrying the chicken hypersensitive site-4 (H54) insulator element. *Mol. Ther.* *15*, 1863–1871.

55. Arumugam, P.I., Urbinati, F., Velu, C.S., Higashimoto, T., Grimes, H.L., and Malik, P. (2009). The 3' region of the chicken hypersensitive site-4 insulator has properties similar to its core and is required for full insulator activity. *PLoS ONE* 4, e6995.
56. Urbinati, F., Arumugam, P., Higashimoto, T., Perumbeti, A., Mitts, K., Xia, P., and Malik, P. (2009). Mechanism of reduction in titers from lentiviral vectors carrying large inserts in the 3'LTR. *Mol. Ther.* 17, 1527–1536.
57. Lee, J.Y., and Lee, H.H. (2018). A new chemical complex can rapidly concentrate lentivirus and significantly enhance gene transduction. *Cytotechnology* 70, 193–201.
58. Shaw, A.M., Joseph, G.L., Jasti, A.C., Sastry-Dent, L., Witting, S., and Cornetta, K. (2017). Differences in vector-genome processing and illegitimate integration of non-integrating lentiviral vectors. *Gene Ther.* 24, 12–20.
59. Frecha, C., Toscano, M.G., Costa, C., Saez-Lara, M.J., Cosset, F.L., Verhoeyen, E., and Martin, F. (2008). Improved lentiviral vectors for Wiskott-Aldrich syndrome gene therapy mimic endogenous expression profiles throughout haematopoiesis. *Gene Ther.* 15, 930–941.
60. Muñoz, P., Toscano, M.G., Real, P.J., Benabdellah, K., Cobo, M., Bueno, C., Ramos-Mejía, V., Menendez, P., Anderson, P., and Martín, F. (2012). Specific marking of HESCs-derived hematopoietic lineage by WAS-promoter driven lentiviral vectors. *PLoS ONE* 7, e39091.
61. Wanisch, K., and Yáñez-Muñoz, R.J. (2009). Integration-deficient lentiviral vectors: a slow coming of age. *Mol. Ther.* 17, 1316–1332.
62. Hanoun, N., Gayral, M., Pointreau, A., Buscail, L., and Cordelier, P. (2016). Initial Characterization of Integrase-Defective Lentiviral Vectors for Pancreatic Cancer Gene Therapy. *Hum. Gene Ther.* 27, 184–192.
63. Sürün, D., Schwäble, J., Tomasovic, A., Ehling, R., Stein, S., Kurrie, N., von Melchner, H., and Schütgen, F. (2018). High Efficiency Gene Correction in Hematopoietic Cells by Donor-Template-Free CRISPR/Cas9 Genome Editing. *Mol. Ther. Nucleic Acids* 10, 1–8.
64. Naldini, L., Blömer, U., Gage, F.H., Trono, D., and Verma, I.M. (1996). Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc. Natl. Acad. Sci. USA* 93, 11382–11388.
65. Loewen, N., Leske, J.A., Chen, Y., Teo, W.L., Saenz, D.T., Peretz, M., Holmes, J.M., and Poeschla, E.M. (2003). Comparison of wild-type and class I Integrase mutant J-FIV vectors in retina demonstrates sustained expression of integrated transgenes in retinal pigment epithelium. *J. Gene Med.* 5, 1009–1017.
66. Park, F., and Kay, M.A. (2001). Modified HIV-1 based lentiviral vectors have an effect on viral transduction efficiency and gene expression in vitro and in vivo. *Mol. Ther.* 4, 164–173.
67. Lieberman, P.M. (2008). Chromatin organization and virus gene expression. *J. Cell. Physiol.* 216, 295–302.
68. Knipe, D.M. (2015). Nuclear sensing of viral DNA, epigenetic regulation of herpes simplex virus infection, and innate immunity. *Virology* 479–480, 153–159.
69. Cara, A., Cereseto, A., Lori, F., and Reitz, M.S., Jr. (1996). HIV-1 protein expression from synthetic circles of DNA mimicking the extrachromosomal forms of viral DNA. *J. Biol. Chem.* 271, 5393–5397.
70. Kilzer, J.M., Stracker, T., Beitzel, B., Meek, K., Weitzman, M., and Bushman, F.D. (2003). Roles of host cell factors in circularization of retroviral dna. *Virology* 314, 460–467.
71. Avettand-Fénoël, V., Hocquedoux, L., Ghosn, J., Cheret, A., Frange, P., Melard, A., Viard, J.P., and Rouzouic, C. (2016). Total HIV-1 DNA, a Marker of Viral Reservoir Dynamics with Clinical Implications. *Clin. Microbiol. Rev.* 29, 859–880.
72. Keaton, M.A., Taylor, C.M., Layer, R.M., and Dutta, A. (2011). Nuclear scaffold attachment sites within ENCODE regions associate with actively transcribed genes. *PLoS ONE* 6, e17912.
73. Cai, S.X., Liu, A.R., He, H.L., Chen, Q.H., Yang, Y., Guo, F.M., Huang, Y.Z., Liu, L., and Qiu, H.B. (2014). Stable genetic alterations of β -catenin and ROR2 regulate the Wnt pathway, affect the fate of MSCs. *J. Cell. Physiol.* 229, 791–800.
74. Kolwí-Shigematsu, T., Maass, K., and Bode, J. (1997). A thymocyte factor SATB1 suppresses transcription of stably integrated matrix-attachment region-linked reporter genes. *Biochemistry* 36, 12005–12010.
75. Feister, H.A., Torungruang, K., Thunyakitpisal, P., Parker, G.E., Rhodes, S.J., and Bidwell, J.P. (2000). NP/NMP4 transcription factors have distinct osteoblast nuclear matrix subdomains. *J. Cell. Biochem.* 79, 506–517.
76. Dunn, K.L., Zhao, H., and Davie, J.R. (2003). The insulator binding protein CTCF associates with the nuclear matrix. *Exp. Cell Res.* 288, 218–223.
77. Zhang, P., Torres, K., Liu, X., Liu, C.G., and Pollock, R.E. (2016). An Overview of Chromatin-Regulating Proteins in Cells. *Curr. Protein Pept. Sci.* 17, 401–410.
78. Deutsch, M.J., Ott, E., Papior, P., and Schepers, A. (2010). The latent origin of replication of Epstein-Barr virus directs viral genomes to active regions of the nucleus. *J. Virol.* 84, 2533–2546.
79. Montes, R., Romero, T., Cabrera, S., Ayllon, V., Lopez-Escamez, J.A., Ramos-Mejía, V., and Real, P.J. (2015). Generation and characterization of the human iPSC line PBMC1-iP54F1 from adult peripheral blood mononuclear cells. *Stem Cell Res. (Amst.)* 15, 614–617.
80. Martín, F., Toscano, M.G., Blundell, M., Frecha, C., Srivastava, G.K., Santamaría, M., Thrasher, A.J., and Molina, I.J. (2005). Lentiviral vectors transcriptionally targeted to hematopoietic cells by WASP gene proximal promoter sequences. *Gene Ther.* 12, 715–723.
81. Demaison, C., Parsley, K., Brouns, G., Scherr, M., Battmer, K., Kinnon, C., Grez, M., and Thrasher, A.J. (2002). High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of immunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. *Hum. Gene Ther.* 13, 803–813.
82. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta Ct) Method. *Methods* 25, 402–408.
83. Torres, R., Garcia, A., Jimenez, M., Rodriguez, S., and Ramirez, J.C. (2014). An integration-defective lentivirus-based resource for site-specific targeting of an edited safe-harbour locus in the human genome. *Gene Ther.* 21, 343–352.
84. Manders, R.J., Koopman, R., Sluijjsmans, W.E., van den Berg, R., Verbeek, K., Saris, W.H., Wagenmakers, A.J., and van Loon, L.J. (2006). Co-ingestion of a protein hydrolysate with or without additional leucine effectively reduces postprandial blood glucose excursions in Type 2 diabetic men. *J. Nutr.* 136, 1294–1299.

RESEARCH ARTICLE

Comparison of Zinc Finger Nucleases Versus CRISPR-Specific Nucleases for Genome Editing of the Wiskott-Aldrich Syndrome Locus

Alejandra Gutierrez-Guerrero,^{1,†} Sabina Sanchez-Hernandez,^{1,†} Giuseppe Galvani,^{1,‡} Javier Pinedo-Gomez,¹ Rocío Martín-Guerra,¹ Almudena Sanchez-Gilabert,¹ Araceli Aguilar-González,¹ Marién Cobo,^{1,2} Philip Gregory,^{3,§} Michael Holmes,³ Karim Benabdellah,^{1,2,**} and Francisco Martín^{1,2,*,***}

¹Centre for Genomics and Oncological Research (GENYO), Pfizer/University of Granada/Andalusian Regional Government, Genomic Medicine Department, Granada, Spain; ²LentiStem Biotech, Granada, Spain; ³Sangamo BioSciences, Point Richmond Tech Center, Richmond, California.

[†]A.G.-G. and S.S.-H. contributed equally to this paper.

[‡]Present address: RCCS San Raffaele Scientific Institute–Diabetes Research Institute (DRI), Milan, Italy.

[§]Present address: Bluebird Bio Inc., Cambridge, Massachusetts.

^{*}K.B. and F.M. share senior authorship.

Primary immunodeficiencies, including Wiskott-Aldrich syndrome (WAS), are a main target for genome-editing strategies using specific nucleases (SNs) because a small number of corrected hematopoietic stem cells could cure patients. In this work, we have designed various WAS gene-specific CRISPR/Cas9 systems and compared their efficiency and specificity with homodimeric and heterodimeric WAS-specific zinc finger nucleases (ZFNs), using K-562 cells as a cellular model and plasmid nucleofection or integration-deficient lentiviral vectors (IDLVs) for delivery. The various CRISPR/Cas9 and ZFN SNs showed similar efficiency when using plasmid nucleofection for delivery. However, dual IDLVs expressing ZFNs were more efficient than dual IDLVs expressing Cas9 and guide RNA or all-in-one IDLVs, expressing Cas9 and guide RNA in the same vector. The specificity of heterodimeric ZFNs and CRISPR/Cas9, measured by increments in γ -H2AX focus formation in WAS-edited cells, was similar for both, and both outperformed homodimeric ZFNs independently of the delivery system used. Interestingly, we show that delivery of SNs, using IDLVs, is more efficient and less genotoxic than plasmid nucleofection. We also show the similar behavior of heterodimeric ZFNs and CRISPR/Cas9 for homology-directed gene knock-in strategies, with 88 and 83% of the donors inserted in the WAS locus, respectively, whereas when using homodimeric ZFNs only 45% of the insertions were on target. In summary, our data indicate that CRISPR/Cas9 and heterodimeric ZFNs are both good alternatives to further develop SN-based gene therapy strategies for WAS. However, IDLV delivery of WAS-specific heterodimeric ZFNs was the best option of all systems compared in this study.

Keywords: genome editing, Wiskott-Aldrich syndrome, integrative deficient lentiviral vectors (IDLV), zinc finger nucleases (ZFN), CRISPR/Cas9, K562

INTRODUCTION

WISKOTT-ALDRICH SYNDROME (WAS) is an X-linked severe primary immunodeficiency characterized by the presence of eczema, thrombocytopenia, recurrent infections, tumor development, and autoimmune diseases.^{1–4} WAS is caused by mutations in the WAS gene, which is localized on the short arm of the X chromosome and encodes a hematopoietic-specific 502-amino acid protein, named WAS pro-

tein (WASP). The only curative treatment for WAS is hematopoietic stem cell (HSC) transplantation with HLA-identical sibling donors or matched unrelated donors. However, compatible donors are not always available and serious side effects due to graft-versus-host disease are also frequent.^{5,6} WAS is therefore a good candidate for gene therapy where gene-modified autologous HSCs expressing the WAS gene are transplanted back into the

*Correspondence: Dr. Francisco Martín, GENYO, Pfizer/University of Granada/Andalusian Regional Government, Genomic Medicine Department, Avenida de la Ilustración 114, 18016 Granada, Spain. E-mail: francisco.martin@genyo.es

patient.⁶ Several clinical trials using gammaretroviral vectors^{7,8} and lentiviral vectors (LVs)^{9,10} have rendered good results in terms of efficiency^{7–10} as well as safety.^{9,10} However, retroviral vectors, including lentiviral vectors, integrate into sites of active transcription and can therefore have undesired side effects on the target cells.^{8,11}

Advances in targeted genome-editing (GE) tools have brought gene therapy the opportunity to specifically modify the affected locus with no or minimal alterations in the target cells.¹² The revolution started with the possibility to generate nucleases that specifically cut in the selected locus in the human genome.^{13–16} These specific nucleases (SNs) recognize specific nucleotide sequences, generating double-strand breaks (DSBs) that the cell repairs by homologous recombination (HR) or by nonhomologous end-joining (NHEJ). Depending on the cell's mechanism used to repair the DSBs and the materials delivered into the cell, we can repair, insert, or delete DNA fragments in the targeted genome. The specificity of SNs can rely on protein–DNA recognition, such as zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), or on RNA–DNA recognition clustered regularly interspaced short palindromic repeats (CRISPR/Cas system).

The CRISPR/Cas system derives from bacterial adaptive immune systems,^{17,18} and it is one of the most powerful GE technologies at the moment.¹⁹ The guide RNA (gRNA) hybridizes with an ~20-nucleotide sequence next to a protospacer-adjacent motif (PAM) and enables the Cas9 endonuclease to cleave the target site, creating DSBs.¹⁶ Various Cas9 nucleases derived from various bacterial species have been used to develop efficient GE tools.¹⁹ Each CRISPR/Cas9 system differs in terms of gRNA design and PAM requirements in order to achieve specific DSBs, and all of them have been shown to efficiently edit the genome; nevertheless, Cas9 from *Streptococcus pyogenes* (SpCas9) has been the most widely used.²⁰ However, CRISPR/Cas9 systems generate DSBs outside of the target sequences (off-target sites) that can differ by several nucleotides from on-target sites and can generate different types of mutations.^{21,22}

ZFNs are chimeric proteins composed of an endonuclease (the bacterial *FokI* restriction enzyme) fused with several zinc finger domains that bind specifically to the target sequence. These zinc fingers consist of 30 amino acids in a $\beta\beta\alpha$ configuration and recognize 3 or 4 base pairs of DNA. Because the cleavage domain must dimerize, the system requires two different ZFNs (left and right) binding to opposite strands of DNA at a certain distance

(5–7 bp).²³ However, it has been reported that the ZFNs can also form homodimers (ZFNleft–ZFNleft and ZFNright–ZFNright), generating DSBs in genome locations different from the intended target (off-target sites).^{24,25} To increase ZFN specificity, Miller and colleagues developed ZFNs harboring a mutated *FokI* enzyme that is active only when forming heterodimers.¹³ The specificity of ZFNs can also be increased by decreasing the intracellular concentration of ZFNs.²⁶

Efficiency and safety are the two main aspects of any therapeutic strategy. Successful GE therapeutic approaches must be able to deliver the SNs (and the donor DNA for gene correction) into a large number of target cells in order to have a clinical benefit. In addition, the delivery system must be innocuous to the target cells. This is a relatively easy task for immortalized cell lines, but primary human cells are more difficult and require special delivery systems. HSCs, the final target for WAS gene therapy, are particularly resistant to the various delivery technologies. Nevertheless, various SNs as well as donor DNA can be efficiently delivered to HSCs by nucleofection,^{27,28} adeno-associated viruses (AAVs),^{29,30} and integration-deficient lentiviral vectors (IDLVs).^{31,32} IDLVs appeared initially to be promising delivery tools for HSC GE because of their low toxicity and high efficiency. However, ZFNs have been delivered efficiently to HSCs, using IDLV systems,^{31,33–35} whereas the delivery of other SNs such as TALENs and CRISPRs has been more difficult.³⁶

The efficiency of SNs can be easily measured by analyzing the percentage of cells harboring insertions and deletions (indels) or incorporating the desired modifications in the targeted sequence. However, safety is more complicated to study, due to the difficulty in detecting off-target cleavage of the SNs. Most studies determine specificity by *in silico* predictions of potential off-target sites followed by experimental analysis with CEL (celery) nuclease (commercialized under the brand name Surveyor) or deep sequencing. However, these biased systems cannot detect unpredicted sites³⁷ as demonstrated by several studies that detect off-target cleavage sites in live cells without any kind of previous restriction.^{21,22,38–40} However, unbiased off-target analysis (GUIDE-seq [genome-wide, unbiased identification of DSBs enabled by sequencing], IDLVs, deep sequencing, etc.) is generally costly and technically challenging. An alternative way to measure safety is to quantify the amount of DSBs generated by the SNs. One of the best markers by which to quantify DSBs is H2AX phosphorylation. H2AX is a variant of H2A that becomes the center of the cellular response when DNA damage

occurs and is phosphorylated, leading to γ -H2AX focus formation.⁴¹ γ -H2AX staining has been used to quantify the DNA damage produced by specific site endonucleases, such as ZFNs.^{13,42} This quantification can be used to analyze the specificity of the various SNs, as the number of foci can be correlated with off-target cleavage.

In this paper we compare the efficiency and safety of various WAS-specific SNs (ZFNs_wt, ZFNs_mt, and CRISPR/Cas9 gRNA9) targeting the same locus and using two different delivery methods (plasmid nucleofection and IDLVs). We use the human leukemia cell line K-562 as target, which is considered to be an adequate model for studying GE because the results can be translated to HSCs,^{27,32} the target for WAS GE therapeutic strategies. We study the efficiency and safety of gene disruption, using plasmid nucleofection or IDLVs to deliver the various SNs. We also use nucleofection of the various SNs together with donor DNA to study the efficiency and safety of GE of the WAS locus by HR. Our data indicate that ZFNs_mt and CRISPR SNs achieve similar efficiencies and specificities by nucleofection, but delivery of ZFNs_mt with IDLVs renders higher efficiencies than with CRISPR/Cas9 in K-562 cells.

MATERIALS AND METHODS

Plasmids

Plasmids encoding original homodimeric ZFNs (ZFNs_wt) targeting intron 1 of the WAS gene (pVAX-N2A-3FN-15724fok and pVAX-N2A-3FN-15755fok) were designed by Sangamo BioSciences (Richmond, CA) as previously described.⁴³ Plasmids encoding optimized ZFNs (ZFNs_mt) (pVAX-N2A-3FN-15724fokEL and pVAX-N2A-3FN-15755fokKK) were described previously⁴⁴ and have a mutated version of the *FokI* nuclease (EL and KK) that increases the affinity for heterodimerization.^{13,45} The ZFN cDNAs were included in the SEWP plasmid⁴⁶ (a self-inactivated LV expressing enhanced green fluorescent protein [eGFP]) under the control of the spleen focus-forming virus (SFFV) promoter) by standard molecular techniques, replacing eGFP for the ZFNs to generate pSZFN24, pSZFN55, pSZFN24EL, and pSZFN55KK.

The Cas9 nuclease cDNA was obtained through gene synthesis (plasmid pUC-coCas9nls2x; GenScript, Piscataway, NJ) of a human codon-optimized cDNA version of the *Streptococcus pyogenes* M1 Cas9 sequence (ref AE004092.2; from nt 854751 to nt 858857) further modified to eliminate the restriction sites for *EcoRI*, *Clal*, *BamHI*, *NotI*, *MluI*, *BmgBI*, *AscI*, *AarI*, *AsiI*, *PstI*, *SbfI*, *XhoI*, *SclI*, *KpnI*,

and *XbaI* and to introduce a *BstXI* site at the +43 position and two *BspI* sites at the +2500 and +2710 positions. In addition, a sequence containing a hemagglutinin (HA) tag and two nuclear localization signals (NLSs) from the simian virus 40 T antigen (SV40T) was included at the 3' end. Finally, the coCas9-HA-nls2x sequence was flanked by a *BamHI*-*AscI* site at the 5' end and a *PstI*-*AarI* site at the 3' end to facilitate cloning into a lentiviral plasmid backbone.

Guide RNAs (gRNAs) were designed by searching for protospacer adjacent motif (PAM) sequences near the cutting site of the WAS-specific ZFNs. Basic Local Alignment Search Tool (BLAST) searches of the various gRNAs came up with only one gRNA (gRNA9: 5'-GAGGCAGGAAGCACCAGGTC-3') with the desired characteristics (targeting the same region as the ZFNs and having no homology with other regions in the genome) (Supplementary Fig. S1; Supplementary Data are available online at www.liebertpub.com/hum). gRNA9 was synthesized by GenScript, together with the scaffold RNA and the U6 promoter, to ensure robust expression of the gRNA molecule (plasmid pU6_gRNA9 containing BbsI-U6-gRNA9-scaffold-BbsI).

SELg9 and SELg9WP plasmids were obtained as follows: a *XhoI*-*XbaI* fragment of the SEWP plasmid containing the 3' long terminal repeat (LTR) was subcloned into a pUC plasmid to obtain pUC-LTR. The gRNA9scaffold was introduced into the 3' LTR *BbsI* site of pUC by standard molecular techniques to obtain pUC-LTR-gRNA9. Finally, a *XhoI*-*XbaI* fragment of pUC-LTR-gRNA9 containing LTR-gRNA9 was used to replace the *XhoI*-*XbaI* fragment of the SE and SEWP lentiviral vectors to obtain the SELg9 and SELg9WP plasmids.

To obtain SCas9Lg9 and SCas9Lg9WP, the eGFP cDNA from SELg9 and SELg9WP were removed by *BamHI* and *PstI* and replaced with a *BamHI*-*PstI* fragment from the pUC-coCas9nls2x plasmid containing the codon-optimized Cas9 cDNA.

To obtain the SCas9Xg9WP and the SCas9Eg9WP plasmids, gRNA9 was amplified by PCR with primers containing the *XhoI* or *EcoRI* sites (see Supplementary Table S1) and inserted into the *XhoI* and *EcoRI* sites of SCas9WP, respectively.

Cell lines and culture media

293T cells (CRL-11268; American Type Culture Collection, Manassas, VA) were grown in Dulbecco's modified Eagle's medium (DMEM; Invitrogen, Edinburgh, Scotland) with GlutaMAX and supplemented with 10% fetal bovine serum (FBS) (PAA Laboratories, Pasching, Austria). The K-562 cell line (chronic myelogenous leukemia) was pur-

chased from the ATCC (CCL-243) and cultured in RPMI 1640 (Gibco-BRL, Middlesex, UK) supplemented with 10% FBS (PAA Laboratories).

Vector production

Lentiviral vectors were produced by cotransfection of 293T cells with three plasmids: (1) a vector plasmid (SEWP, SCas9WP, SELg9W, SCas9Lg9, SCas9Lg9W, SCas9Xg9W, SCas9Eg9W, SZFN24W, SZFN55W, SZFN24ELW, or SZFN55KKW), (2) the human immunodeficiency virus (HIV) packaging plasmid pCMvdr8.91 for integration-competent LVs or pCMvdr8.91D64V for IDLVs, and (3) the envelope plasmid pMD2.G.⁴⁷ Vector production was performed as previously described.⁴⁸ Briefly, 293T cells were plated on amine-coated petri dishes (Sarstedt, Newton, NC), in order to ensure 80–90% confluence for transfection. The three plasmids were resuspended in LipoD293 (SigmaGen, Gaithersburg, MD) as described previously.⁴⁹ The plasmid–LipoD293 mixture was added to pre-washed cells and incubated for 6–8 hr, when new medium was added. After 48 hr viral supernatants were collected, filtered through a 0.45- μ m (pore size) filter (Nalgene, Rochester, NY), and concentrated either by ultrafiltration at 2000 \times g at 4°C, using 100KD centrifugal filter devices (Amicon Ultra-15; Millipore, Billerica, MA) or by ultracentrifugation (23,000 rpm, 2 hr, at 4°C; Beckman Coulter, Brea, CA). Concentrated vectors were either directly used or aliquot and frozen at –80°C.

Viral titers of eGFP-expressing LVs and IDLVs were determined by transduction of 1×10^5 K-562 cells per well in 48-well tissue culture plates (BD Biosciences, San Jose, CA), using a 4-fold serial dilution of the supernatants and then assaying the percentage of eGFP-positive cells by fluorescence-activated cell-sorting (FACS) analysis 5 days (LVs) or 48 hr (IDLVs) after transduction. LV and IDLV particle numbers per milliliter were determined with a lentivirus titration kit (qPCR lentivirus titration kit; Applied Biological Materials [ABM], Richmond, BC, Canada) according to the manufacturer's protocol. Briefly, samples were diluted 1:100 and 1:1000 in 100 μ l of phosphate-buffered saline (PBS) before lysis. Viral lysis was performed with 2 μ l of diluted viral supernatant added to 18 μ l of the virus lysis buffer provided. The qRT-PCR program was performed with an Applied Biosystems 7300 real-time PCR system (Thermo Fisher Scientific, Waltham, MA), using the following qRT-PCR parameters: 1 cycle of reverse transcription (42°C for 20 min), 1 cycle of enzymatic activation (95°C for 10 min), 40 cycles of denaturation (95°C for 15 sec), and 40 cycles of annealing/extension (60°C for 1 min).

Cell transduction

293T and K-562 cells were washed with Dulbecco's PBS (1 \times) (PAA Laboratories) and dissociated with Invitrogen TrypLE Express (Thermo Fisher Scientific) diluted one-third in PBS (293T cells) or directly centrifuged at 1250 rpm for 5 min and counted. Cells (1×10^5) were plated in 24- or 48-well plates (BD Biosciences) and incubated for 5 hr with the various LVs or IDLVs. Serial dilutions were made to determine titers. Medium was changed after 5 hr.

Cell nucleofection

For gene disruption, K-562 cells (1×10^6) were nucleofected with the various pairs of ZFN-expressing plasmids (3 μ g each; SZFN24W/pSZFN55W [ZFNs_wt] or SZFN24ELW/SZFN55KKW [ZFNs_mt]), with two plasmids expressing Cas9 and gRNA9 (SCas9WP/SELg9W), with 6 μ g of the all-in-one CRISPR plasmid SCas9Lg9W, or with 6 μ g of control (irrelevant) plasmid (S δ EWP, expressing mutated eGFP). For the homology directed (HD) gene insertion assay, the donor plasmid (12–18 μ g) was also included together with the selected SN plasmids. Nucleofection was performed with an Amaxa Nucleofector II and solution V (Lonza, Basel, Switzerland), applying program T16 and following the nucleofection protocol for K-562 cells.

Flow cytometry

Controls and transduced K-562 cells ($1\text{--}4 \times 10^5$) were collected and washed with cold FACS buffer (PBS containing 2% FBS and 0.5% bovine serum albumin [BSA]). 293T cells were dissociated with TrypLE Express, washed with PBS containing 2% FBS, and centrifuged at 1250 rpm for 5 min and washed with cold FACS buffer. Cells were stained with 7-aminoactinomycin D (7-AAD) viability dye (eBioscience, San Diego, CA) and viable cells (7-AAD-negative cells) were gated for the subsequent analysis. Finally, cells were analyzed for GFP expression, using a FACSCanto II flow cytometer (Becton Dickinson, Franklin Lakes, NJ) equipped with FACSDiva analysis software (BD Biosciences).

Quantification of cleavage efficiency of WAS target site

Genomic DNA extraction from K-562 cells was performed with a QIAamp DNA mini kit (Qiagen, Chatsworth, CA), in accordance with the protocol for cells in suspension. The genomic regions flanking the CRISPR and ZFNs target site for each treatment were amplified by PCR using hWASP5Fw/hWASP5Rev primers (primers are listed in Supplementary Table S1). The 2-kb fragments were purified with a QIAquick spin

column (Qiagen), in accordance with the manufacturer's protocols. Total purified PCR product (400 ng) was mixed with 2 μ l of 10 \times *Taq* DNA polymerase PCR buffer and ultrapure water to a final volume of 20 μ l, and subjected to a re-annealing process to enable heteroduplex formation. After reannealing, products were treated with Surveyor nuclease (Integrated DNA Technologies, San Diego, CA), in accordance with the manufacturer's recommended protocols, and analyzed on 2% agarose gels. Two fragments of 1.2 and 0.8 kb will appear if the SNs have generated a DSB at the target site. The percentage of cleavage was determined by densitometry of the 2.0-, 1.2-, and 0.8-kb bands, using the following formula: $\text{indel (\%)} = 100 \times (1 - \sqrt{(1 - fcut)})$, where $fcut = (a+b)/(a+b+c)$, where a is integrated intensity of uncleaved DNA while b and c are integrated intensity of cleavage products after background correction.⁵⁰

DSB analysis: analysis of H2AX phosphorylation

K-562 cells were nucleofected or transfected with the various SNs and, 48 hr later, 2.5×10^5 cells were collected and washed with cold FACS buffer (PBS containing 2% FBS and 0.5% BSA). Cells were then fixed (80% methanol for 5 min), permeabilized (0.1% PBS-Tween for 20 min), and incubated with PBS-10% goat serum for 10 min to block nonspecific protein-protein interaction, followed by incubation with anti-histone γ -H2AX (phosphorylated at Ser-139) antibody (diluted 1:100; Abcam, Cambridge, UK) for 30 min at room temperature. Subsequent to washing with PBS, cells were incubated with goat anti-rabbit IgG H+L Alexa Fluor 647 antibody (diluted 1:10,000; Abcam) for 30 min at room temperature. Isotype control antibody was rabbit IgG monoclonal (Abcam) (0.1 μ g), also incubated for 30 min at room temperature. After washing with PBS, cells were analyzed for γ -H2AX expression, using the FACSCanto II flow cytometer (Becton Dickinson) equipped with FACSDiva analysis software (BD Biosciences).

Measurement of γ -H2AX by image cytometry. K-562 cells were nucleofected or transfected with the various SNs as described above. Formation of γ -H2AX in response to DSBs was analyzed by image cytometry, using a FlowCollect histone H2AX phosphorylation assay kit (Merck Millipore) according to the manufacturer's instructions. γ -H2AX expression was measured with an ImageStream X Mark II imaging flow cytometer (Merck Millipore) and analyzed with IDEAS software, using the Spot Wizard. To determine the number of

γ -H2AX foci in each cell, the Peak Mask was applied, which gives the number of foci per nucleus according to an internal algorithm. The number of individual masks in a cell was enumerated with the Spot Count feature and plotted for every cell in histograms.

Off-target detection

Possible off-target sites for the CRISPR/Cas9_gRNA system were identified through BLAST⁵¹ and Bowtie⁵² algorithms, using the entire sequence of gRNA9 (GAGGCAGGAAGGACCAGGTCTGG) or just the first 16 bp including the PAM sequence (GAAGGACCAGGTCTGG). For ZFNs we used the genome-wide tag scanner for nuclease off-sites,⁵³ using the binding sites of both ZFNs as target. We selected the eight best potential off-targets for each nuclease (see Supplementary Tables S2 and S3) and analyzed indel generation by Surveyor as described previously and using the primer pairs described in Supplementary Tables S2 and S3.

Efficiency and fidelity of HD donor insertion into the WAS locus

K-562 cells were nucleofected with plasmids expressing the various SNs (pSZFN24/pSZFN55, pSZFN24EL/pSZFN55KK, or pSCas9WP/SELg9) and donor DNA expressing the eGFP and neomycin resistance (NeoR) genes. Cells transfected with the nucleases and donor DNA were selected with neomycin for 20 days, and GFP-positive cells were sorted with a FACSAria II (BD Biosciences) and seeded into a 96-well plate to obtain single cells (1 cell/5 wells). Clones were grown in the presence of neomycin, and genomic DNA was extracted from more than 50 clones from each condition. The DNAs were used as templates for PCRs, using (1) a primer pair (hWASP5Fw/hWASP5Rev) flanking the WAS target sequence, outside the donor homology domains; (2) two primer pairs that amplify the endogenous WAS and donor DNA at the 5' junction (hWASP5Fw/eGFPRev) and 3' junction (eGFPFw/hWASP5Rev); and (3) a primer pair (SFFV1aFw/NeoRRev) that amplifies an internal fragment from the donor DNA (see Supplementary Fig. S2 and Supplementary Table S1). All the PCRs were carried out with Kappa *Taq* polymerase (Kapa Biosystems, Wilmington, MA).

Statistical analysis

All data are expressed as means \pm SEM. Statistical comparisons were performed by Student *t* test, with the assumption of normal distribution. Statistical significance was defined as $p < 0.05$.

RESULTS

Construction of WAS-specific ZFNs and CRISPR/Cas9 systems

WAS-specific homodimeric ZFNs (ZFNs_wt) and heterodimeric ZFNs (ZFNs_mt) were designed by Sangamo as described in Materials and Methods. A WAS-specific CRISPR/Cas9 system was designed and home-made, based on previously published systems.^{16,54} The Cas9 nuclease cDNA was obtained through gene synthesis (GenScript) of a codon-optimized cDNA version of the *Streptococcus pyogenes* M1 Cas9 sequence incorporating an HA tag and two NLSs from SV40T at the 5' end (see Materials and Methods for details). We searched for WAS-specific guide RNAs (gRNAs) targeting the same sequence as the ZFNs and without potential off-target sites (see Materials and Methods for details and Supplementary Fig. S1). Only one gRNA (gRNA9) fulfilled these requirements and was synthesized by GenScript together with the scaffold and the U6 promoter.

The efficiency and safety of specific nucleases can vary depending on expression levels, which are highly influenced by the delivery system used. We have therefore constructed various plasmids (Fig. 1; see Materials and Methods for details) that can be used to deliver the SNs by plasmid nucleofection and by IDLVs. Each ZFN (ZFN24, ZFN55, ZFN24EL, and ZFN55KK) is expressed by a plasmid harboring the SFFV promoter and all the elements required to be packed into lentiviral particles. The same plasmids are used to express

Cas9 and gRNA9 separately. K-562 cells were nucleofected with the plasmids or transduced with IDLV particles, and the various samples were analyzed for the expression of ZFNs, Cas9, and gRNA 2 days later to verify expression of the various SNs and gRNA (data not shown).

Delivery of SNs by nucleofection showed similar efficiency and specificity of ZFNs and CRISPR/Cas9

We first compared the cutting efficiency and specificity of both ZFNs versus CRISPR/Cas9 systems, using nucleofection as the delivery method. Plasmids expressing a ZFNs_wt pair (SZFN24W, SZFN55W), a ZFNs_mt pair (SZFN24ELW and SZFN55KKW), and dual CRISPR (Scas9W and SELg9W) were nucleofected into K-562 cells, using the Amaxa Nucleofector, and analyzed for specific (Surveyor) and unspecific (γ -H2AX-staining) DSB generation (Fig. 2). Although highly variable, we observed similar WAS gene-targeting efficiencies of all systems (Fig. 2A, right). Nonspecific DSB generation by the various SNs was analyzed first by measuring the increments in γ -H2AX staining by flow cytometry (FACS) and using K-562 cells nucleofected with irrelevant plasmid as controls (Mock) (Fig. 2A). The CRISPR/Cas9_gRNA9 and ZFNs_mt systems showed similar γ -H2AX staining compared with K-562 cells nucleofected with irrelevant plasmid, indicative of the low off-target activity of these SNs (Fig. 2A). However, at similar cutting efficiencies, ZFNs_wt presented two to four

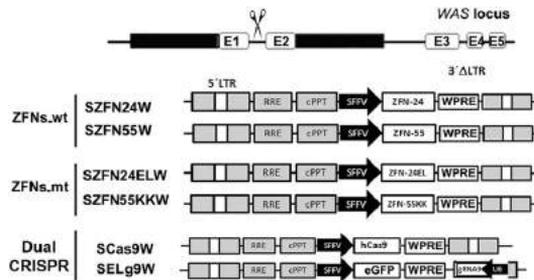
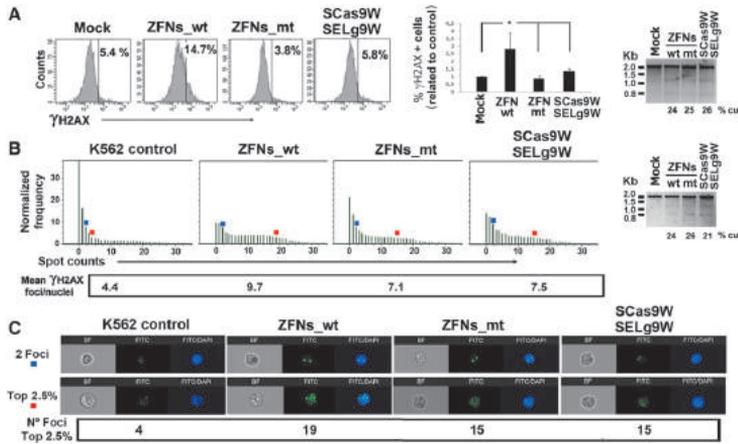


Figure 1. Plasmid constructs used to express the various WAS-specific nucleases. *Top:* A drawing of the WAS target site. Each zinc finger nuclease (ZFN) pair (ZFNs_wt and ZFNs_mt) is expressed by two lentiviral vector plasmids expressing each ZFN (24 and 55 or 24EL and 55KK) through the spleen focus-forming virus (SFFV) promoter and containing the posttranscriptional regulatory element of woodchuck hepatitis virus (WPRE). For comparison, the CRISPR system (Cas9/gRNA, *bottom*) is also expressed by two different plasmids: one expressing human codon-optimized Cas9 cDNA, using the same backbone as the ZFNs (SCas9W), and the other expressing gRNA9 through the U6 promoter and eGFP through the SFFV promoter (SELg9W).



Downloaded by Duke Univ from online.liebertpub.com at 11/02/17. For personal use only.

Figure 2. Comparison of WAS-specific zinc finger nucleases and CRISPR systems delivered by nucleofection. **(A)** Quantification of double-strand breaks (DSBs) by fluorescence-activated cell-sorting (FACS) analysis. *Left:* Representative histograms for K-562 cells nucleofected with irrelevant plasmid (Mock) and ZFNs_wt (SZFN24W and SZFN55W), ZFNs_mt (SZFN24ELW and SZFN55KW), or dual CRISPR (SCas9W and SELg9W) plasmids, stained with antibodies against γ -H2AX to detect DSBs. The percentage of cells positive for γ -H2AX is indicated for each histogram. *Middle:* Graph showing increment in DSBs (γ -H2AX-positive cells) of K-562 cells nucleofected with the various plasmids expressing ZFNs_wt, ZFNs_mt, or dual CRISPR related to Mock K-562. Averages of at least two separate experiments are shown ($^*p < 0.05$; two-tailed unpaired Student *t* test). *Right:* Analysis of the cutting efficiency of the various specific nucleases. Nuclease activity at the target site was monitored by PCR, digestion with the Surveyor nuclease, and detection by agarose gel electrophoresis. The cutting efficiencies are shown as percentages of cleavage (estimated by densitometry of the bands) and indicated under each lane. **(B)** γ -H2AX foci quantification by ImageStream cytometry. *Left:* Histograms showing the frequency of cells (normalized frequency) that contain from 0 to 35 γ -H2AX foci per nucleus (spot counts) in K-562 control cells (nucleofected without plasmid) and cells nucleofected with plasmids expressing ZFNs_wt, ZFNs_mt, or dual CRISPR systems. Populations harboring two foci (blue squares) and populations having a frequency of 2.5% (red squares) are indicated in each sample. *Right:* Analysis of the cutting efficiency of the various specific nucleases (SNs) in the experiment analyzed with the ImageStream cytometer. Nuclease activity at the target site was monitored as mentioned above. The cutting efficiencies are shown as percentages of cleavage and indicated under each lane. **(C)** Representative images of K-562 control cells (nucleofected without plasmid) and K-562 cells nucleofected with plasmids expressing ZFNs_wt, ZFNs_mt, or dual CRISPR SNs harboring two foci per nucleus (top; blue square) and of the populations that have a normalized frequency of 2.5% (top 2.5%; bottom, red square). The numbers of foci that correspond to the top 2.5% for each sample are indicated at the bottom.

times higher percentages of γ -H2AX-positive cells compared with Mock (Fig. 2A).

We next investigated whether γ -H2AX staining on FACS correlates with γ -H2AX focus formation (indicative of DSB repair) and whether plasmid nucleofection could also affect DSB generation on K-562 cells. K-562 cells were therefore nucleofected with ZFNs_wt, ZFNs_mt, CRISPR, irrelevant plasmids, or without plasmid DNA. Forty-eight hours later the various samples were analyzed for γ -H2AX focus formation, using the ImageStream and IDEAS analysis software (Fig. 2B and C) (see Materials and Methods for details). This system allowed us to verify that γ -H2AX staining indeed identified γ -H2AX spots (see Fig. 2B and C, blue and red squares). The IDEA

software was used to quantify the amount of spots per nucleus, plotted in histograms, and the mean γ -H2AX foci per nucleus calculated in each sample (Fig. 2B, left). The results showed that, as observed by FACS, K-562 cells nucleofected with ZFNs_mt, CRISPR, and irrelevant plasmids had similar DSB levels when measured as mean γ -H2AX foci per nucleus (Fig. 2B: 7.1 and 7.5; and data not shown), whereas ZFNs_wt rendered the highest DSB levels (9.7). We also observed the same number of foci in populations that have a normalized frequency of 2.5% (top 2.5%) in K-562 cells nucleofected with the ZFNs_mt or CRISPR plasmids (Fig. 2B and C, red squares; 15 foci per nucleus), whereas K-562 cells nucleofected with the ZFNs_wt plasmid showed up to

19 foci per nucleus. Of note, K-562 cells nucleofected with no DNA showed lower DSB levels than K-562 cells nucleofected with ZFNs_mt, CRISPR, or irrelevant plasmid (Fig. 2B and C, K-562 control vs. ZFNs_mt and CRISPR; and data not shown).

All together, these data showed that nucleofection with plasmids expressing ZFNs_mt and CRISPR had similar low off-target activity that was below the limit of detection by γ -H2AX staining, in part due to the genotoxicity of plasmid nucleofection. However, nucleofection with plasmids expressing ZFNs_wt increased DSB generation compared with nucleofection with irrelevant plasmid and must therefore be due to the off-target activity of ZFNs_wt.

We also analyzed the specificity of the various nucleases by *in silico* analysis of potential off-target sites. For CRISPR/Cas9_gRNA9, we used BLAST⁵¹ and Bowtie⁵² algorithms to search for potential gRNA9 off-target sites. For ZFNs we used the genome-wide tag scanner for nuclease off-sites.⁵³ We selected the eight highest scoring off-targets for each nuclease (see Supplementary Tables S2 and S3) and analyzed indel generation by Surveyor ($\approx 2\%$ sensitivity). However, no detectable off-target activity was observed for any SNs, using this technology (data not shown), probably due to the low sensitivity of this technique.

IDLVs-ZFNs are more efficient than IDLVs-CRISPR but have similar specificity

Delivery of SNs by IDLVs is an interesting alternative for SN delivery because of their wide tropism, their ability to incorporate various elements (SNs, donor, miRNA), and the potential to be

used *in vivo*. We therefore generated IDLVs expressing each of the CRISPR components (Cas9 and gRNA9) and each of the ZFNs_mt, using the plasmids showed in Fig. 1. K-562 cells were transduced with equivalent particle numbers per cell (PNC, 400) of each IDLV pair (SCas9W/SELg9W or SZFN24ELW/SZFN55KKW) and the cutting efficiencies and mRNA expression levels were determined (Fig. 3). Our data showed higher cutting efficacies of IDLVs-ZFNs_mt (35.6%) compared with the dual IDLVs-CRISPR system (25.8%) (Fig. 3A); this could be due, at least in part, to the lower mRNA expression levels of the latter (Fig. 3B). Indeed, our analysis showed that IDLVs-CRISPR expressed 3–10 times lower mRNA levels compared with IDLVs-ZFNs_mt, probably due to a longer messenger (Cas9, 4.2 kb vs. ZFNs, 1.1 kb).

To increase the efficiency of the CRISPR system, we constructed various all-in-one IDLVs to express both Cas9 and gRNA9 (Fig. 4A, top) and analyzed their efficiency in K-562 cells, using similar volumes (Fig. 4B, bottom) or PNC (Supplementary Fig. S3B). Optimized IDLVs achieved cutting efficacies ranging from 21.2% to 35.5% (PNC, 250–350), showing that these all-in-one IDLVs-CRISPR can be used for WAS GE. However, although the CRISPR all-in-one IDLVs outperformed the dual IDLVs, their efficacy was still two to four times lower compared with the dual IDLVs-ZFNs (Fig. 4B and Supplementary Figs. S3B and S4). Interestingly, WAS gene disruption using any of the IDLV systems did not cause any harm to the treated cells (Supplementary Fig. S5). Because the SCas9Lg9W IDLVs showed the best efficiencies we used this construct for further studies.

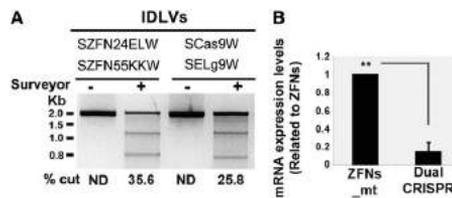


Figure 3. Comparative analysis of the efficiency of IDLVs ZFNs_mt and dual CRISPR systems to edit the WAS locus. (A) Surveyor analysis of K-562 cells transduced with integration-deficient LVs (IDLVs) expressing ZFNs_mt (SZFN24ELW and SZFN55KKW, left lanes) or dual CRISPR (SCas9W and SELg9W, right lanes) at a particle number per cell (PNC) of 400 each. Nuclease activity at the target site was monitored by PCR, digestion with the Surveyor nuclease (+), and detection by agarose gel electrophoresis. The activities are shown as the percentage of cleavage (estimated by densitometry of the bands) and indicated under each lane. (B) mRNA quantification of K-562 cells transduced with ZFNs_mt or dual CRISPR IDLVs. Quantification of the total amount of mRNA was performed as described in Materials and Methods, using common primers to the WPRE region of the IDLVs and normalized to ZFNs_mt. Data represent means \pm SEM of at least three different experiments (** $p < 0.01$; two-tailed unpaired Student *t* test).

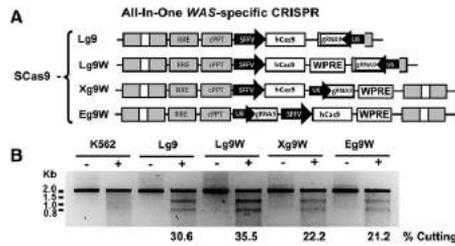


Figure 4. Development of all-in-one IDLVs expressing Cas9 and gRNAs. **(A)** Drawing of the various all-in-one lentiviral plasmid constructs expressing the human codon-optimized *Cas9* gene through the SFV promoter and gRNA9 through the U6 promoter. The U6-gRNA9 expression cassette was inserted inside the 3' LTR (Lg9 and Lg9W), 3' of the *Cas9* gene (Xg9W) or 5' of the *Cas9* gene (Eg9W). The WPRE was inserted in all constructs with the exception of Lg9 (top) for comparison with Lg9W. **(B)** Analysis of the cutting efficiency of the various all-in-one IDLVs, using 200 μ l of unconcentrated supernatants on 100,000 K-562 cells (PNC, 250–350). Nuclease activity at the target site was monitored by PCR (+), digestion with the Surveyor nuclease (+), and detection by agarose gel electrophoresis. The activities are shown as percentages of cleavage (estimated by densitometry of the bands) and indicated under each lane.

Downloaded by Duke Univ from online.liebertpub.com at 11/02/17. For personal use only.

We next generated SCas9Lg9W IDLV particles as well as IDLVs expressing each ZFN and compared their efficiency and specificity to disrupt the *WAS* locus. It is important to note that the specificity of the various SNs must be compared in target cells edited at similar levels because higher cutting efficiencies at the *WAS* locus will result in higher γ -H2AX staining and could also render higher off-target activity. Therefore, K-562 cells were cotransduced with IDLV-ZFN24W/IDLV-ZFN55W (IDLVs-ZFNs_wt), IDLV-ZFN24ELW/IDLV-ZFN55KKW (IDLVs-ZFNs_mt), or all-in-one SCas9Lg9W IDLVs at various PNCs in order to reach similar cutting efficiencies. Forty-eight hours later we analyzed DSBs by γ -H2AX and indel generation at the target site (*WAS* locus) by Surveyor (Fig. 5A, right and Supplementary Fig. S6). We required two to four times higher PNCs for IDLVs-CRISPR compared with IDLVs-ZFNs in order to reach similar efficiencies. Interestingly, as observed with nucleofection, FACS analysis showed similar levels of γ -H2AX-positive cells in CRISPR- and ZFNs_mt-treated K-562 cells, whereas cells treated with ZFNs_wt showed higher γ -H2AX staining, indicating higher off-target activity (Fig. 5A and Supplementary Fig. S6).

We further compared efficacy and specificity of the dual IDLVs-ZFNs and the all-in-one IDLVs-CRISPR by incubating 10^6 K-562 cells with equal volumes of concentrated IDLVs and analyzing the cutting efficiency and DSB generation (Fig. 5B and C). As observed previously, we detected lower efficiency of the all-in-one IDLVs-CRISPR compared with the dual IDLVs-ZFNs_mt or IDLVs_wt (Fig. 5B, right; 22% vs. 32 and 36%, respectively). We also confirmed that

γ -H2AX staining and FACS correlated (at least partially) with the amount of γ -H2AX foci observed with the ImageStream cytometer (Fig. 5B, left). Indeed, K-562 cells transduced with IDLVs-ZFNs_wt showed a mean number of γ -H2AX foci per nucleus higher than that of K-562 cells transduced with IDLVs-ZFNs_mt (Fig. 5B, left; mean γ -H2AX foci per nucleus of 5.5 vs. 3.1, respectively). We cannot compare ZFNs_mt with CRISPR in this experiment because the *WAS* cutting efficacy was lower in the latter (32 vs. 22%). However, we could show that the average amounts of DSBs in control cells (NT; non-transduced) were similar to those found in K-562 cells edited with IDLVs-ZFNs_mt and IDLVs-CRISPR (3.4, 3.1, and 3.9, respectively; Fig. 5B). In the same direction, we also observed a similar number of foci in populations that have a normalized frequency of 1% (top 1%) in K-562 control cells and in K-562 cells transduced with IDLVs-ZFNs_mt or IDLVs-CRISPR (Fig. 5B and C, red squares; 17, 15, and 17 foci per nucleus, respectively), whereas K-562 cell transduced with ZFNs_wt IDLVs showed up to 24 foci per nucleus. These data indicate that delivery of SNs by IDLV had a low impact on the target cells, something that was also observed in Fig. 3.

Interestingly, we could also observe that IDLVs showed lower γ -H2AX focus formation compared with nucleofection in spite of achieving higher cutting efficacies of the *WAS* gene (compare Fig. 2B with Fig. 5B and Supplementary Fig. S7). Indeed, K-562 *WAS* GE using IDLVs-ZFNs_mt and IDLVs-CRISPR achieved 32 and 22% cuts and mean γ -H2AX foci per nucleus of 3.1 and 3.9, respectively, whereas the same cells edited using nucleofection

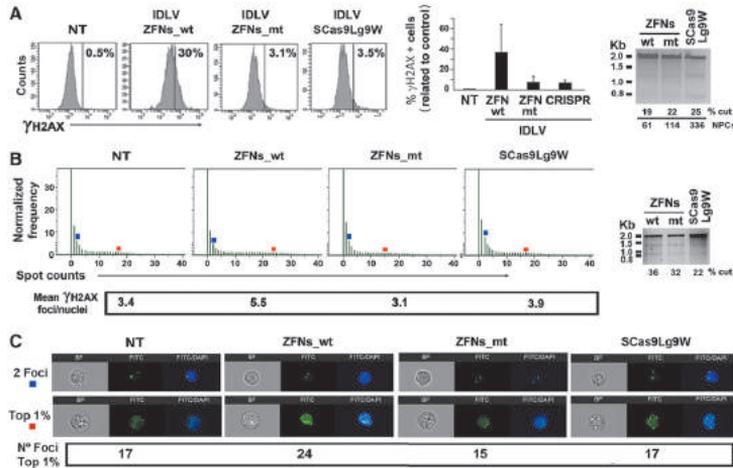


Figure 5. Comparison of WAS-specific zinc finger nucleases and all-in-one CRISPR system delivered by IDLVs. **(A)** Quantification of DSBs by FACS analysis. *Left:* Representative histograms for human control K-562 cells (NT), K-562 cells transduced with IDLVs expressing each homodimeric ZFN (IDLV-ZFNs_wt pair), each heterodimeric ZFN (IDLV-ZFNs_mt pair), or a unique IDLV expressing Cas9 and gRNA9 (IDLV-SCas9Lg9W) stained with antibodies against γ -H2AX to detect double-strand breaks. The percentages of cells positive for γ -H2AX are indicated for each histogram. *Middle:* Graph showing increments in double-strand breaks (γ -H2AX-positive cells) related to untreated K-562 cells (NT) or K-562 cells transduced with the various IDLV combinations reaching 15–30% cutting efficiency. Averages of at least two separate experiments are indicated. *Right:* Analysis of the cutting efficiency of the various SNs, using different NPCs. Nuclease activity at the target site was monitored by PCR, digestion with the Surveyor nuclease, and detection by agarose gel electrophoresis. The activities are shown as percentages of cleavage (estimated by densitometry of the bands) and are indicated under each lane. **(B)** γ -H2AX focus quantification by ImageStream cytometry. *Left:* Histograms showing the frequency of cells (normalized frequency) that contain from 0 to 40 γ -H2AX foci per nucleus (spot counts) in control cells (NT, nontransduced) and cells transduced (10^6) with $100 \mu\text{l}$ of IDLV (concentrated $30 \times$) ZFNs_wt, ZFNs_mt, or SCas9Lg9W. Populations harboring two foci (blue squares) and populations having a frequency of 1% (red squares) are indicated in each sample. *Right:* Analysis of the cutting efficiency of the various IDLVs in the experiment analyzed with the ImageStream cytometer. Nuclease activity at the target site was monitored as mentioned above. The activity is shown as the percentage of cleavage and is indicated under each lane. **(C)** Images of K-562 cell controls (NT) and K-562 cells transduced with IDLVs-ZFNs_wt, IDLVs-ZFNs_mt, or IDLVs-SCas9Lg9W harboring two foci per nucleus (*top*, blue square) and representative images of the populations that have a normalized frequency of 1% (*top 1%*, *bottom*, red square). The number of foci that correspond to the top 1% for each sample are indicated at the bottom.

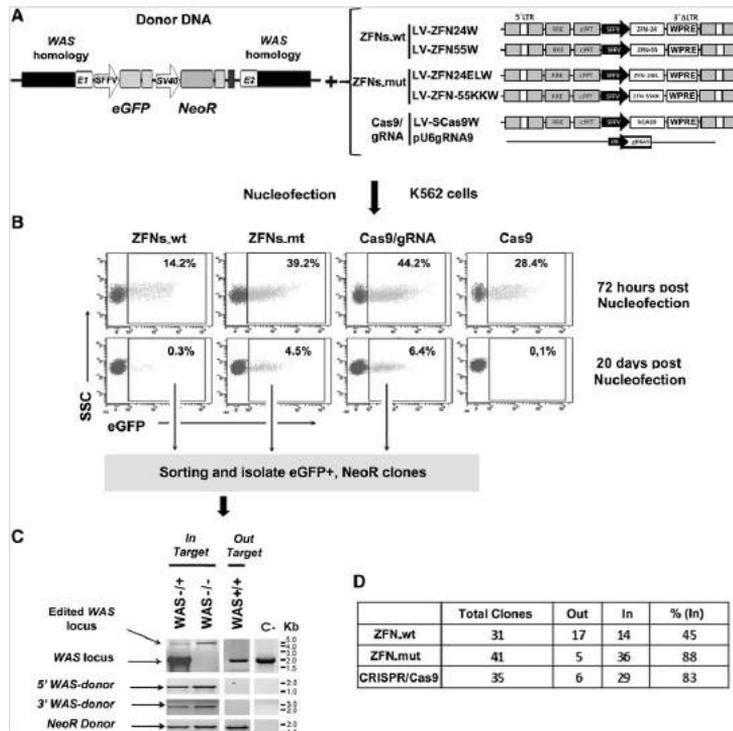
of plasmid ZFNs_mt and CRISPR achieved efficiencies of 26 and 21% cuts and mean γ -H2AX foci per nucleus of 7.1 and 7.5, respectively.

All together, these data indicate that the WAS-specific ZFNs_mt have similar specificity compared with CRISPR-gRNA9, and that delivery of SNs by IDLVs has lower impact on DSB generation in target cells compared with nucleofection of plasmid DNA.

Efficiency and specificity of ZFNs and CRISPR for HD GE of WAS locus

We next investigated the efficiency and specificity of WAS-specific ZFNs_wt, ZFNs_mt, and CRISPR systems to insert donor DNA into the WAS

locus. We designed a donor containing eGFP and neomycin resistance gene expression cassettes flanked by 1-kb homologous arms to the WAS locus (Fig. 6A). Plasmids expressing the various SNs were nucleofected together with the donor DNA, and eGFP expression was analyzed on days 3 and 20 (Fig. 6B). We achieved transfection of about 30–45% of the K-562 cells (Fig. 6B, top plots), which renders between 0.3 and 6% of K-562 cells stably expressing eGFP (Fig. 6B, bottom plots) when co-transfected with the SNs. After 20 days, most of these cells stably expressing eGFP have integrated the donor DNA into the chromosome. To study only those cells in which donor integration has occurred,



Downloaded by Duke Univ from online.liebertpub.com at 11/02/17. For personal use only.

Figure 6. Analysis of the efficiency and safety of ZFNs and CRISPR/Cas9 for homologous recombination (HR)-directed WAS genome editing. **(A)** Scheme showing the procedure to insert a donor DNA (harboring eGFP and neomycin resistance [NeoR] expression cassettes) into the WAS locus, using various specific nucleases. **(B)** Plots indicating the GFP percentage 72 hr and 20 days after nucleofection with donor DNA and ZFNs_wt (left), ZFNs_mt (second left), the CRISPR/Cas9 system containing gRNA (second right), and the CRISPR/Cas9 system without gRNA (right). After 20 days, eGFP⁺ cells were sorted, and more than 100 clones were grown in medium with neomycin for further analysis. **(C)** Analysis of the integration pattern of GFP⁺ clones. Electrophoresis shows an example of the analysis performed to determine in-target and out-target integrations of the donor DNA in the various clones. **(D)** Table showing the specificity of donor integration with the various SNs. The table shows the total number of clones analyzed, the number of clones containing donor integration outside the WAS locus (Out), and the number of clones harboring donor integration at the WAS locus (In), as well as the percentage of clones harboring donor insertion at the correct sites.

we sorted the eGFP⁺ cells and generated clones in the presence of neomycin. These clones (GFP⁺NeoR⁺) must have integrated the complete donor DNA. The various clones were analyzed by PCR as depicted in Fig. 6C to differentiate between donor integrations in

the WAS locus (In-target) and integrations outside the WAS locus (Out-target) (see also Supplementary Fig. S2 for details). The results showed that ZFNs_mt and CRISPR achieved 88 and 83% of the insertions in the targeted locus, respectively (Fig. 6C

and D). On the other hand, cells transfected with donor DNA and ZFNs_wt had 55% of the insertions landing outside the WAS locus (Fig. 6D).

These data are in line with the cutting specificity data of the various SNs obtained previously. To further confirm the integration of donor DNA in the WAS locus, we analyzed WASP expression levels in homozygous and heterozygous clones by FACS analysis. Supplementary Figure S8 shows the absence or reduction of WASP expression in homozygous and heterozygous clones, respectively.

DISCUSSION

In this study we have compared the efficiency and safety of various WAS-specific nucleases (ZFNs_wt, ZFNs_mt, and CRISPR/Cas9_gRNA9) for gene disruption as well as for HD gene insertion into the WAS locus, using K-562 cells as a cellular model. We used K-562 cells as target because they are considered an adequate model for studying GE in hematopoietic cells and the results can be translated to HSCs,^{27,32} the final target for WAS gene therapy. Because the efficiency and safety of SNs depend on the efficiency of the delivery system as well as the expression levels achieved on target cells, we generated a set of plasmids that can be used for delivery of ZFNs and the CRISPR/Cas9 system either by nucleofection or IDLVs.

We first showed that the method used to deliver the SNs affected differently the efficiency of ZFNs and CRISPR/Cas9 systems. Indeed, both SNs showed similar WAS cleavage when delivered by plasmid nucleofection. However, delivery of the CRISPR system using two IDLVs showed lower cutting activity compared with ZFNs, even when the CRISPR system was delivered as all-in-one IDLVs. Our data also indicate that the lower efficiency of the IDLVs-CRISPR system was in part due to low expression levels of Cas9 mRNA. Indeed, we have hypothesized that the Cas9 levels achieved by the IDLVs-CRISPR system must be close to the threshold required to achieve GE. This could explain why IDLVs-CRISPR systems require higher NPCs than IDLVs-ZFNs systems and also the observed differences in efficacy of various experiments using IDLVs-CRISPR, even when similar NPCs are used. This hypothesis is partially supported by the fact that integrative LVs-CRISPR are always more robust and achieve similar or even higher cutting efficacies than integrative LV-ZFNs (integrative LVs express higher mRNA levels than their IDLV counterparts) (data not shown).

Our desire to improve IDLV delivery of the CRISPR system prompted us to design IDLVs ex-

pressing both Cas9 and gRNA in the same vector and to optimize their expression levels by inserting the gRNA-expressing cassette into the LTR and including the WPRE. The all-in-one IDLVs-WAS-specific CRISPR systems showed good WAS gene cleavage, although efficacy was still two to four times lower compared with the dual system expressing each ZFN under two different IDLVs. In addition, we could not detect any differences in cell viability when comparing similar cutting efficiencies, indicating that IDLVs-ZFNs could be a better option than IDLVs-CRISPR-gRNA9.

In a clinical setting it is fundamental that SNs cut only in the target site because off-target cleavage could generate a wide set of mutations that could, in the worst scenario, cause cellular transformation. We therefore analyzed the specificities of the WAS-targeted homodimeric and heterodimeric ZFNs versus CRISPR/Cas9_gRNA9 by Surveyor analysis of potential off-target sites (see Materials and Methods and data not shown) as well as for γ -H2AX staining. γ -H2AX immunostaining allows for identification of DSB foci that can be generated by SNs at the target locus but also at the off-target loci, and can therefore be used to measure them.^{13,41,42} However, DSBs are also generated in the DNA of cells growing under normal conditions, and this noise can sometimes interfere with the off-target analysis. In addition, higher cutting efficiencies of the SNs will also render higher γ -H2AX staining, and therefore this methodology can be used only to compare DSB generation in cells edited at similar levels. Indeed, we verified in our system that higher cutting efficacies at the WAS locus rendered higher γ -H2AX signal (Supplementary Fig. S6). Another technical difficulty in off-target assessment based on the γ -H2AX signal is uncertainty concerning whether it correlates with focus formation due to DSB generation. Because FACS cannot discriminate between signals coming from spots (DSBs) or from uniform nuclear staining (unspecific), we proceeded to analyze off-target generation using γ -H2AX staining and ImageStream cytometry, which allows the identification, localization, and quantification of DSB foci. Using this technology we showed that, independent of the delivery method used, heterodimeric ZFNs and CRISPR/Cas9-gRNA9 behaved similarly and both showed lower levels of nonspecific DSB generation compared with homodimeric ZFNs. These data indicate that the CRISPR/Cas9 system can be designed to reach similar specificities compared with heterodimeric ZFNs. Interestingly, we also showed that IDLVs showed lower γ -H2AX focus formation compared with plasmid nucleofec-

tion in spite of achieving even higher levels of WAS gene disruption. In this direction, we also observed that control plasmid nucleofection increased the amount of DSBs in target cells compared with nucleofection without DNA or with untreated cells. Therefore, we must take into account the nonspecific DSBs generated by nucleofection rather than by the off-target activity of the SNs.

Another important aspect to take into account before translating gene repair strategies into the clinic is the fidelity of GE. Indeed, although SNs enhanced the efficiency and fidelity of HD recombination,^{55,56} the administered donor DNAs can be integrated into DSBs generated by SN off-targets, cellular stress, or by the presence of homologous microdomains.^{57,58} These undesired insertions can be highly deleterious for the targeted cells. We therefore analyzed whether the efficiency and fidelity of donor insertion into the WAS locus was affected by the SN used. Our system allowed the identification and isolation of those K-562 clones having the donor integrated in their genome, facilitating the analysis of the integration sites on individual cells. PCR analysis of the various clones showed again a similar behavior of heterodimeric ZFNs and CRISPR/Cas9_gRNA9, with 88 and 83% of the donors inserted in the WAS locus, respectively. Interestingly, when using homodimeric ZFNs only 45% of the insertions landed in the WAS locus, which corroborates the lower specificity of these ZFNs compared with heterodimeric ZFNs and with CRISPR/Cas9_gRNA. These data are in line with previously published data analyzing targeted integration of donor DNAs.³² However, considering that DSBs can be generated by the plasmid nucleofection process, a good proportion of the donor insertions outside of the WAS locus must be due to the delivery process. However, the differences observed between ZFNs_mt and CRISPR versus ZFNs_wt must be due to the higher off-target activity of the latter.

Our studies have been performed in K-562 cells, which are easy to grow and with similar resistance

for gene modification compared with many primary human cells, including HSCs.²⁷ The possibility to generate clones for K-562 cells allowed us to investigate the specificity of HD gene insertion in single cells. K-562 cells have been widely used as hematopoietic cellular models for GE of a wide range of target genes including WAS,⁴⁴ AAVS1,^{59–61} ILR2, the HBB locus,⁶² CCR5, ABCC11, OPN1SW, CH3L1, CEL, VEGFA, and CABPB.⁶³ Results observed in K-562 cells can be translated to HSCs, although the efficiency of HD gene repair is much lower in HSCs compared with K-562 cells.³²

Taken together, our data indicate that both CRISPR/Cas9_gRNA9 and heterodimeric ZFNs rendered good efficiencies and specificities for WAS GE and that IDLVs outperformed plasmid nucleofection as delivery vehicle. We also showed that ZFNs delivered by IDLVs render higher efficiencies than CRISPR-IDLVs, probably due to lower expression levels.

ACKNOWLEDGMENTS

This work has been financed by Fondo de Investigaciones Sanitarias ISCIII (Spain) and Fondo Europeo de Desarrollo Regional (FEDER) from the European Union, through research grants P112/01097, P115/02015 (F.M.), and ISCIII Red de Terapia Celular (TerCel: RD12/0019/0006) (F.M.); by the CICE and CS de la Junta de Andalucía FEDER/Fondo de Cohesion Europeo (FSE) de Andalucía 2007–2013 through research grants P09-CTS-04532, PI-57069, PI-0001/2009, and PAIDI-Bio-326 (F.M.); PI-0160/2012, PI-0014-2016 (K.B.), and PI-0407/2012, PI-0318/2014 (M.C.). A.G.-G. and S.S.-H. are doctoral students from the Biomedicine Doctorate Program from Granada University (Spain). S.S.-H. is funded by Plan de Empleo Juvenil (MINECO).

AUTHOR DISCLOSURE

All of the authors declare no competing interests and no personal financial interest.

REFERENCES

- Massaad, M.J., Ramesh N., Geha R.S. Wiskott-Aldrich syndrome: a comprehensive review. *Ann N Y Acad Sci* 2013;1285:26–43.
- Kirchhausen T. Wiskott-Aldrich syndrome: a gene, a multifunctional protein and the beginnings of an explanation. *Mol Med Today* 1998;4:300–304.
- Jin Y, Mazza C, Christie JR, et al. Mutations of the Wiskott-Aldrich syndrome protein (WASP): hotspots, effect on transcription, and translation and phenotype/genotype correlation. *Blood* 2004;104:4010–4019.
- Ochs HD. The Wiskott-Aldrich syndrome. *Springer Semin Immunopathol* 1998;19:435–458.
- Moratto D, Gilliani S, Bonfim C, et al. Long-term outcome and lineage-specific chimerism in 194 patients with Wiskott-Aldrich syndrome treated by hematopoietic cell transplantation in the period 1980–2009: an international collaborative study. *Blood* 2011;118:1675–1684.

6. Worth AJ, Thrasher AJ. Current and emerging treatment options for Wiskott-Aldrich syndrome. *Expert Rev Clin Immunol* 2015;11:1015–1032.
7. Braun CJ, Bortz K, Paruzynski A, et al. Gene therapy for Wiskott-Aldrich syndrome: long-term efficacy and genotoxicity. *Sci Transl Med* 2014;6:227ra33.
8. Bortz K, Schmidt M, Schwarzer A, et al. Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *N Engl J Med* 2010;363:1918–1927.
9. Hachein-Bey Abina S, Gaspar HB, Blondeau J, et al. Outcomes following gene therapy in patients with severe Wiskott-Aldrich syndrome. *JAMA* 2015;313:1550–1563.
10. Aiuti A, Biasco L, Scaramuzza S, et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* 2013;341:1233–1235.
11. Cesana D, Ranzani M, Volpin M, et al. Uncovering and dissecting the genotoxicity of self-inactivating lentiviral vectors *in vivo*. *Mol Ther* 2014;22:774–785.
12. Porteus M. Genome editing: a new approach to human therapeutics. *Annu Rev Pharmacol Toxicol* 2015;56:163–190.
13. Miller JC, Holmes MC, Wang J, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* 2007;25:779–785.
14. Redondo P, Prieto J, Munoz IG, et al. Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature* 2008;456:107–111.
15. Christian M, Cermak T, Doyle EL, et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 2010;186:757–761.
16. Jinek M, Chylinski K, Fontana I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012;337:816–821.
17. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, et al. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* 2005;60:174–182.
18. Barrangou R, Fremaux C, Deveau H, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007;315:1709–1712.
19. Wright AV, Nunez JK, Doudna JA. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* 2016;164:29–44.
20. Doudna JA, Charpentier E. Genome editing: the new frontier of genome engineering with CRISPR-Cas9. *Science* 2014;346:1258096.
21. Tsai SQ, Zheng Z, Nguyen NT, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 2015;33:187–197.
22. Wang X, Wang Y, Wu X, et al. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol* 2015;33:175–178.
23. Durai S, Mani M, Kandavelou K, et al. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res* 2005;33:5978–5990.
24. Szczepek M, Brondani V, Buchel J, et al. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol* 2007;25:786–793.
25. Cornu TI, Thibodeau-Beganny S, Guhl E, et al. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol Ther* 2008;16:352–358.
26. Pattanayak V, Ramirez CL, Joung JK, et al. Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. *Nat Methods* 2011;8:765–770.
27. Mandal PK, Ferreira LM, Collins R, et al. Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell* 2014;15:643–652.
28. Ye L, Wang J, Tan Y, et al. Genome editing using CRISPR-Cas9 to create the HPHF genotype in HSPCs: an approach for treating sickle cell disease and β -thalassaemia. *Proc Natl Acad Sci U S A* 2016;113:10661–10665.
29. Sather BD, Romano Ibarra GS, Sommer K, et al. Efficient modification of CCR5 in primary human hematopoietic cells using a megaTAL nuclease and AAV donor template. *Sci Transl Med* 2015;7:307ra156.
30. Ling C, Bhukhai K, Yin Z, et al. High-efficiency transduction of primary human hematopoietic stem/progenitor cells by AAV6 vectors: strategies for overcoming donor-variation and implications in genome editing. *Sci Rep* 2016;6:35495.
31. Joglekar AV, Hollis RP, Kufnec G, et al. Integrase-defective lentiviral vectors as a delivery platform for targeted modification of adenosine deaminase locus. *Mol Ther* 2013;21:1705–1717.
32. Genovese P, Schirolli G, Escobar G, et al. Targeted genome editing in human repopulating hematopoietic stem cells. *Nature* 2014;510:235–240.
33. Lombardo A, Genovese P, Beausejour CM, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol* 2007;25:1298–1306.
34. Cornu TI, Cathomen T. Targeted genome modifications using integrase-deficient lentiviral vectors. *Mol Ther* 2007;15:2107–2113.
35. Maggio I, Goncalves MA. Genome editing at the crossroads of delivery, specificity, and fidelity. *Trends Biotechnol* 2015;33:280–291.
36. Gori JL, Hsu PD, Maeder ML, et al. Delivery and specificity of CRISPR-Cas9 genome editing technologies for human gene therapy. *Hum Gene Ther* 2015;26:443–451.
37. Martin F, Sanchez-Hernandez S, Gutierrez-Guerrero A, et al. Biased and unbiased methods for the detection of off-target cleavage by CRISPR/Cas9: an overview. *Int J Mol Sci* 2016;17(9).
38. Cencic R, Miura H, Malina A, et al. Protospacer adjacent motif (PAM)-distal sequences engage CRISPR-Cas9 DNA target cleavage. *PLoS One* 2014;9:e109213.
39. Kucsu C, Arslan S, Singh R, et al. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol* 2014;32:677–683.
40. Gabriel R, Lombardo A, Arens A, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol* 2011;29:816–823.
41. Kinzer A, Wu W, Staudt C, et al. γ -H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. *Nucleic Acids Res* 2008;36:5678–5694.
42. Olsen PA, Solhag A, Booth JA, et al. Cellular responses to targeted genomic sequence modification using single-stranded oligonucleotides and zinc-finger nucleases. *DNA Repair (Amst)* 2009;8:298–308.
43. Moehle EA, Rock JM, Lee YL, et al. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc Natl Acad Sci U S A* 2007;104:3055–3060.
44. Toscano MG, Anderson P, Munoz P, et al. Use of zinc-finger nucleases to knock out the WAS gene in K562 cells: a human cellular model for Wiskott-Aldrich syndrome. *Dis Model Mech* 2013;6:544–554.
45. Doyon Y, Vo TD, Mendel MC, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods* 2011;8:74–79.
46. Demaison C, Parsley K, Brouns G, et al. High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [corection of immunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. *Hum Gene Ther* 2002;13:803–813.
47. Toscano MG, Frasca C, Ortega C, et al. Efficient lentiviral transduction of Hepesvirus saimiri immortalized T cells as a model for gene therapy in primary immunodeficiencies. *Gene Ther* 2004;11(12):956–961.
48. Benabdellah K, Cobo M, Munoz P, et al. Development of an all-in-one lentiviral vector system based on the original TetR for the easy generation of Tet-ON cell lines. *PLoS One* 2011;6:e23734.
49. Benabdellah K, Gutierrez-Guerrero A, Cobo M, et al. A chimeric HS4-SAR insulator (IS2) that prevents silencing and enhances expression of lentiviral vectors in pluripotent stem cells. *PLoS One* 2014;9:e84268.
50. Ran FA, Hsu PD, Lin CY, et al. Double nicking by RNA-guided CRISPR-Cas9 for enhanced genome editing specificity. *Cell* 2013;154:1380–1389.
51. Altschul SF, Gish W, Miller W, et al. Basic Local Alignment Search Tool. *J Mol Biol* 1990;215:403–410.

52. Langmead B, Trapnell C, Pop M, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;10:R25.
53. Cradick TJ, Ambrosini G, Iseli C, et al. ZFN-site searches genomes for zinc finger nuclease target sites and off-target sites. *BMC Bioinformatics* 2011;12:152.
54. Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013;339:819–823.
55. Porteus MH, Carroll D. Gene targeting using zinc finger nucleases. *Nat Biotechnol* 2005;23:967–973.
56. Urnov F.D., Rebar E.J., Holmes M.C., et al. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* 2010;11:636–646.
57. Holkers M, Maggio I, Henriques SF, et al. Adenoviral vector DNA for accurate genome editing with engineered nucleases. *Nat Methods* 2014; 11:1051–1057.
58. Corni TI, Cathomen T. Quantification of zinc finger nuclease-associated toxicity. *Methods Mol Biol* 2010;649:237–245.
59. Orlando SJ, Santiago Y, DeKelver RC, et al. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res* 2010; 38:e152.
60. Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science* 2013;339:823–826.
61. Ma Y, Ramezani A, Lewis R, et al. High-level sustained transgene expression in human embryonic stem cells using lentiviral vectors. *Stem Cells* 2003;21:111–117.
62. Cottle RN, Lee CM, Bao G. Treating hemoglobinopathies using gene-correction approaches: promises and challenges. *Hum Genet* 2016;135:993–1010.
63. Kim S, Kim D, Cho SW, et al. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* 2014;24:1012–1019.

Received for publication March 23, 2017;
accepted after revision September 17, 2017.

Published online: September 18, 2017.