UNIVERSIDAD AUTÓNOMA DE MADRID
Facultad de Ciencias
Departamento de Biología Molecular

Targeted gene therapy in a mouse model of Fanconi anemia

María José del Pino del Barrio

DOCTORAL THESIS
2016

UNIVERSIDAD AUTONOMA DE MADRID

FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

Targeted gene therapy in a mouse model of Fanconi anemia

Memoria presentada por MARÍA JOSÉ DEL PINO DEL BARRIO licenciada en Biotecnología, para optar al grado de doctor por la Universidad Autónoma de Madrid con mención Internacional.

Directores de Tesis

Susana Navarro Ordoñez

Juan Antonio Bueren Roncero

Madrid, 2016

CENTRO DE INVESTIGACIONES ENERGÉTICAS, MEDIOAMBIENTALES Y TECNOLÓGICAS (CIEMAT)-
CENTRO DE INVESTIGACIONES BIOMÉDICAS EN RED DE ENFERMEDADES RARAS (CIBERER-ISCIII)-
INSTITUTO DE INVESTIGACIÓN SANITARIA DE LA FUNDACIÓN JIMÉNEZ DÍAZ (IIS-FJD/UAM)

Susana Navarro Ordoñez, investigadora postdoctoral de la División de Terapias Innovadoras del Sistema Hematopoyético del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, y **Juan Antonio Bueren Roncero**, jefe de dicha unidad, certifican que la memoria adjunta titulada ***Targeted gene therapy in a mouse model of Fanconi anemia*** ha sido realizada por María José del Pino del Barrio bajo la dirección conjunta de los que suscriben, y cumple las condiciones exigidas para optar al título de Doctor con mención Internacional por la Universidad Autónoma de Madrid.

Susana Navarro Ordoñez

Juan Antonio Bueren Roncero

El trabajo de investigación descrito en esta memoria ha sido realizado en la División de Terapias Innovadoras del Sistema Hematopoyético del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), el Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER-ISCIII) y la Unidad Mixta de Terapias Avanzadas CIEMAT/Instituto de Investigación Sanitaria de la Fundación Jiménez Díaz (IIS-FJD/UAM).

Para su ejecución, el trabajo de investigación realizado ha contado con la colaboración de los siguientes Programas de Investigación:

- Séptimo Programa Marco de la Comisión Europea (Proyecto PERSIST; Ref 222878).
- Ministerio de Economía y Competitividad (Proyecto SAF 2012-39834).
- Fondo de Investigaciones Sanitarias, Instituto de Salud Carlos III (RETICS-RD06/0010/0015; RD12/0019/0023).
- Dirección General de Investigación de la Comunidad de Madrid (Proyecto CellCAM; Ref S2012/BMD-2420).
- Programa de transferencia de tecnología en el campo de la terapia génica de la Fundación Botín.

María José del Pino del Barrio ha disfrutado de una beca-contrato para Personal Investigador en Formación (FPI) del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT) del Ministerio de Economía y Competitividad (BOE nº137, de 9 de junio de 2011), y de un contrato del Instituto de Investigación Sanitaria de la Fundación Jiménez Díaz (IIS-FJD/UAM).

A mis padres y a mi hermana,

a Miguel,

por todo su esfuerzo, y el amor que me han dado.

Le Petit Prince s'en fut revoir les roses:

- Vous êtes belles, mais vous êtes vides, leur dit-il encore. On ne peut pas mourir pour vous. Bien sûr, ma rose à moi, un passant ordinaire croirait qu'elle vous ressemble. Mais à elle seule elle est plus importante que vous toutes, puisque c'est elle que j'ai arrosée. Puisque c'est elle que j'ai mise sous globe. Puisque c'est elle que j'ai abritée par le paravent. Puisque c'est elle dont j'ai tué les chenilles (sauf les deux ou trois pour les papillons). Puisque c'est elle que j'ai écoutée se plaindre, ou se vanter, ou même quelquefois se taire. Puisque c'est ma rose.

Et il revint vers le renard:

- Adieu, dit-il...

- Adieu, dit le renard. Voici mon secret. Il est très simple: on ne voit bien qu'avec le cœur. L'essentiel est invisible pour les yeux.

- L'essentiel est invisible pour les yeux, répéta le petit prince, afin de se souvenir.

- C'est le temps que tu as perdu pour ta rose qui fait ta rose si importante.

- C'est le temps que j'ai perdu pour ma rose...- fit le Petit Prince, afin de se souvenir.

- Les hommes ont oublié cette vérité, dit le renard. Mais tu ne dois pas l'oublier. Tu deviens responsable pour toujours de ce que tu as apprivoisé. Tu es responsable de ta rose...

- Je suis responsable de ma rose... - répéta le petit prince, afin de se souvenir.

Chapitre XXI, Le Petit Prince - Antoine de St Exupéry

El Principito fue a ver nuevamente a las rosas:

- Sois bellas, pero estáis vacías -les dijo de nuevo-. No se puede morir por vosotras. Seguramente un transeúnte cualquiera creería que mi rosa se os parece. Pero ella sola es más importante que todas vosotras, puesto que es ella la rosa a quien he regado. Puesto que es ella la rosa a quien he abrigado bajo un globo. Puesto que es ella a quien protegí con un biombo. Puesto que es ella la rosa cuyas orugas maté (salvo dos o tres que se hicieron mariposas). Puesto que es ella la rosa a quien escuché quejarse, o alabarse, o incluso, algunas veces, callarse. Puesto que es mí rosa.

Y volvió hacia el zorro:

- Adiós – dijo.

- Adiós – dijo el zorro. – Aquí está mi secreto. Es muy simple: sólo se ve bien con el corazón. Lo esencial es invisible a los ojos.

- Lo esencial es invisible a los ojos – repitió el principito a fin de acordarse.

- Es el tiempo que perdiste por tu rosa lo que hace a tu rosa tan importante.

- Es el tiempo que perdí por mi rosa... – dijo el principito a fin de acordarse.

- Los hombres han olvidado esta verdad – dijo el zorro. – Pero tú no debes olvidarla. Eres responsable para siempre de lo que has domesticado. Eres responsable de tu rosa...

- Soy responsable de mi rosa... - repitió el principito a fin acordarse.

Capítulo XXI, El Principito - Antoine de St Exupéry

INDEX

INDEX

I. ABBREVIATIONS/ABREVIATURAS

A

A	Adenine
AAV	Adeno-Associated Virus
AAVS1	Adeno-Associated Virus Site 1
ADA-SCID	Adenosine Deaminase Severe Combined Immunodeficiency
ADN	Ácido Desoxirribonucleico
AdV	Adenovirus
AE	Adverse Effects
AF	Anemia de Fanconi
ALD	Adrenoleukodystrophy
AML	Acute Myeloid Leukemia
APC-Cy	Allophycocyanin-Cyanine
ATM	Ataxia Telangiectasia Mutated protein
ATR	Ataxia Telangiectasia and RAD3-related protein
a.u.	Arbitrary Units

B

β-geo	β-galactosidase-neo fusion gene
B-NHEJ	Back-up Non-Homologous End Joining
β-Thal	β-thalassemia
BACH1	BTB domain And CNC Homolog 1 protein
BLM	Bloom-Syndrome Recq Like helicase
BM	Bone Marrow
BMF	Bone Marrow Failure
bp	Pair of Bases
BRCA1/2	Breast Cancer type 1 or 2 proteins
BRIP1	BRCA1-Interacting Protein 1

C

C	Cytosine
C-Kit	Receptor tyrosine kinase c-Kit
Cas9	CRISPR Associated protein 9
CCR5	C-C Chemokine Receptor type 5
CDC42	Cell Division Control protein 42 homolog
CFCs	Colony Forming Cells
CFUs	Colony Forming Units
CHK1/2	Checkpoint Kinase 1 or 2
CMHs	Células Madre Hematopoyéticas
CMPHs	Células Madre y Progenitores Hematopoyéticos
CMV	Cytomegalovirus
CRISPR	Cluster Regularly Interspaced Short Palindromic Repeats
Ct	Cycling Threshold
CtIP	CtBP-Interacting protein
CXCR4	Chemoquine Receptor 4

D

d	Days
D-loop	Displacement loop
D-NHEJ	DNA-PKcs dependent Non-Homologous End Joining
DAPI	4’,6-Diamidino-2-Phenylindole
DDT	DNA Damage Tolerance
DEB	Diepoxybutane
dHJ	Double Holliday Junction
DMEM	Dulbecco's Modified Eagle’s Medium
DNA	Deoxyribonucleic Acid
DNA2	DNA replication helicase/nuclease 2

DNA-PKcs	DNA-dependent Protein Kinase
dpt	Days Post-Transplant
DSB	Double Strand Break
DSBR	Double Strand Break Repair
DUB	Deubiquitinase

E

EGFP	Enhance Green Fluorescent Protein
EME1	Essential Meiotic Structure-Specific Endonuclease1
EPO	Erythropoietin
ER	Enzyme Replacement
ERCC1	Excision Repair Cross-Complementation group 1 protein
ERCC4	Excision Repair Cross-Complementation group 4 protein
ESCs	Embryonic Stem Cells
ESGCT	European Society of Gene and Cell Therapy

F

FA	Fanconi Anemia
FA-A	Fanconi anemia complementation group A (or any of the other groups -B/ C/ D1/ D2/ E/ F/ G/ I/ J/ L/ M/ N/ O/ P/ Q/ R/ S/ T/ V)
FAAP	Fanconi Anemia Associated Proteins
FAN1	FANCD2/FANCI-Associated Nuclease 1
FANCA/FANCA	Fanconi anemia gene or protein A (or any of the other groups of Fanconi anemia genes or proteins above-mentioned)
FACS	Fluorescence-Activated Cell Sorting
FITC	Fluorescein IsoTiocyanate
FISH	Fluorescence *In Situ* Hybridization
Fw	Forward

G

G	Guanine
G-CSF	Granulocyte Colony-Stimulating Factor
GG-NER	Global Genome Nucleotide Excision Repair
gDNA	Genomic DNA
GMP	Good Manufacturing Practice
γRV	Gamma Retrovirus
GSH	Genomic *Safe Harbor*
GT	Gene Therapy
GvHD	Graft-versus-Host Disease

H

h	Hours
H2AX	H2A histone family member X
HA	Hemagglutinin
Hb	Hemoglobin
HD	Healthy Donor
HDR	Homology-Directed Repair
hFlt3	Human FMS-Like Tyrosine kinase 3 Ligand
hGCSF	Human Granulocyte Colony Stimulating Factor
hIL-6	Human Interleukin-6
hIL-11	Human Interleukin-11
HIV-1	Human Immunodeficiency Virus 1
HLA	Human Leukocyte Antigen
HPV	Human Papillomavirus
HR	Homologous Recombination
HRR	Homologous Recombination mediated Repair
HSCs	Hematopoietic Stem Cells

HSCT	Hematopoietic Stem Cell Transplantation
HSPCs	Hematopoietic Stem and Progenitor Cells
hTPO	human Thrombopoietin
HygR	Hygromycin B Resistance gene

I

ICL	Interstrand Cross-Links
IDLV	Integrase-Deficient Lentiviral Vector
IL-1β	Interleukin 1 Beta
IMDM	Iscove's Modified Dulbecco's Medium
INDELS	Insertions and Deletions
INF-γ	Interferon Gamma
iPSCs	Induced Pluripotent Stem Cells
IR	Ionizing Radiation
IVF-PGD	*In Vitro* Fertilization and Preimplantation Genetic Diagnosis

K

kb	Kilo-base/s

L

L	Left TALEN monomer
LacZ	β-galactosidase gene
LCL	Lymphoblastic Cell Line
LIG	Ligase
Lin^-	Lineage negative cells
LSK	Lineage negative, Sca-1 positive, c-Kit positive cells
LTR	Long Terminal Repeat
LV	Lentivirus

M

m-IL3	Mouse Interleukin-3
*MBS85/*MBS85	Myosin Binding Subunit 85 locus/protein
MDS	Myelodysplasic Syndrome
MEFs	Mouse Embryonic Fibroblasts
mFGF2	Mouse Fibroblast Growth Factor Basic
MFI	Mean Fluorescence Intensity
mHA	Mouse Homology Arm
MHF	MHF1-MHF2-FAAP24 complex
MHF1/2	Histone-Fold complex subunit 1/2
MLD	Metachromatic Leukodistrophy
MMC	Mitomycin C/ Mitomicina C
MN	Meganuclease
MRN	MRE11-RAD50-NBS1 complex
mRNA	Messenger Ribonucleotid Acid
mSCF	Mouse Stem Cell Factor
mTALEN	Mouse Transcription Activator-Like Effector Nucleases
MUS81	Structure-Specific Endonuclease Subunit

N

N/A	Not Analysed
NBS	Nijmegen Breakage Syndrome 1 protein
NeoR	Neomycin Resistance gene
NER	Nucleotide Excision Repair
NHEJ	Non-Homologous End Joining
NK	Natural Killer cells
NK	NK-mTALEN pair
NN	NN-mTALEN pair

NLS	Nuclear Localization Signal
NR	Not Reported
nt	Nucleotide/s

P

P53BP1	Tumor Protein P53 Binding Protein 1
PALB2	Partner And Localizer of BRCA2 protein
PAM	Protospacer-Adjacent Motif
PARI	PCNA-Associated Recombination Inhibitor
PARP-1	Poly (AND-Ribose) Polymerase 1
PB	Peripheral Blood
PBS	Phosphate-Buffered Saline
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase Chain Reaction
PE	Phycoerythrin
PGK	Phosphoglycerate Kinase 1
PLT	Platelets
pmol	Picomoles
Pol	Polymerase
*PPP1R12C/*PPP1R12	Protein Phosphatase 1, Regulatory Subunit 12C locus/protein
PuroR	Puromycin Resistance Gene
P/S	Penicillin/Streptomycin

Q

qPCR	Quantitative Polymerase Chain Reaction

R

R	Right TALEN monomer
rAAV	Recombinant Adeno-Associated Virus

RAD6/18	RAD homologs 6 or 18
RAD51	RAD51 recombinase homolog
RAD52	RAD52 recombinase homolog
RAD51C	RAD51 paralog C
RNAi	Interference Ribonucleic Acid
ROS	Reactive Oxygen Species
RPA	Replication Protein A
RT	Room Temperature
RVD	Repeat Variable Di-residues
Rw	Reverse
R.D.	Real Decreto

S

Sca-1	Stem Cells Antigen-1
SCC	Squamous Cell Carcinomas
SCD	Sickle Cell Disease
SCID	Severe Combined Immunodeficiency
SCID-X1	X-Linked Severe Combined Immunodeficiency
SDS	Sodium Dodecyl Sulfate
SDSA	Synthesis-Dependent Strand Annealing
sgRNA	Single Guide Ribonucleotid Acid
SIN γ-RV	Self-Inactivating Gamma Retroviral vector
SIN-LV	Self-Inactivating Lentiviral vector
SLX1	Structure-Specific Endonuclease subunit 1
SLX4	Structure-Specific Endonuclease subunit 4
SNM1A/B	Suppressor of Nuclear Mitochondrial endoribonuclease 1 subunit A or B protein
ssDNA	Single-Stranded DNA

ssODNs	Single-Stranded Oligonucleotides
S.D.	Standard Deviation

T

T	Thymine
T-ALL	T cell Acute Lymphoblastic Leukemia
TC-NER	Transcription-Coupled Nucleotide Excision Repair
TALE	Transcription Activator-Like Effector
TALEN	Transcription Activator-Like Effector Nuclease
TG	Terapia Génica
TI	Targeted Integration
TLS	Translesion Synthesis
TNF-α	Tumor Necrosis Factor Alfa
TRS/RBS	Terminal Resolution Site/ Rep Binding Site

U

UAF1	USP1-Associated Factor 1
UBE2T	Ubiquitin conjugating Enzyme E2 T
UBZ	Ubiquitin-Binding Zinc Finger domain
USP1	Ubiquitin Specific Peptidase 1
UTR	Untranslated Region
UV	Ultraviolet Radiation

W

W	Week/s
WAS	Wiskott-Aldrich Syndrome
WB	Western Blot
WRN	Werner syndrome, RecQ helicase-like
WT	Wild Type

X

X-CGD	X-linked Chronic Granulomatous Disease
X-RCC4	X-Ray Repair Cross Complementing 4
X-RCC9	X-Ray Repair Complementing Defective in Chinese Hamster Cells 9
XP	Xeroderma Pigmentosum
XPA	Xeroderma Pigmentosum A-complementing protein
XPC	Xeroderma Pigmentosum C-complementing protein
XPF	Xeroderma Pigmentosum F-complementing protein

Z

ZF	Zinc Finger
ZFN	Zinc Finger Nuclease

Note: Some of the names of genes or proteins indicated in the thesis in parentheses are explained in detail in abbreviations but not in text due to format requirements.

II. SUMMARY

SUMMARY

Fanconi anemia (FA) is an inherited disease associated with bone marrow failure (BMF) and cancer predisposition. This disease is caused by mutations in any of the 20 *FANC* genes that belong to a DNA repair pathway known as FA/BRCA pathway. Gene therapy (GT) approaches with autologous hematopoietic stem and progenitor cells (HSPCs) may constitute an efficient treatment for FA patients. In fact, conventional non-targeted GT trials based on the use of lentiviral vectors are currently under development. Because the risk of insertional oncogenesis can not be completely ruled out in any of the GT trials that are at present undergoing with integrative vectors, gene-targeting approaches have been proposed as safer alternatives. These strategies are based on the use of artificial nucleases that generate double strand breaks at specific locations in the genome which can be repaired by the homologous recombination pathway of the cells. This pathway has been exploited to facilitate the integration, into specific sites of the cell genome, of donor templates harboring the gene of interest flanked by two homology arms.

In this study we have conducted a gene-editing strategy aiming at the insertion of reporter and therapeutic donors into the mouse *Mbs85* locus, an ortholog of the human *AAVS1 safe harbor* locus. In order to achieve our goal, TALE nucleases (TALEN) were nucleofected together with different donor templates carrying either the human therapeutic *FANCA* gene or the *EGFP* (Enhance Green Fluorescent Protein) reporter gene, both of which were under the regulation of a phosphoglycerate kinase (PGK) promoter. This strategy of gene editing was implemented by means of nucleofection of the TALEN and donor as DNA plasmids in *Fanca*$^{-/-}$ (FA-A) mouse embryonic fibroblasts (MEFs) and also in mouse HSPCs (mHSPCs) from WT and FA-A mice.

We have first demonstrated the cleavage activity of designed TALEN in the *Mbs85* locus of both FA-A MEFs and mHSPCs from WT and FA-A mice. Targeted integration (TI), as well as evidence of hFANCA expression were shown in gene-edited FA-A MEFs. Moreover, evidence of phenotypic correction was demonstrated in these cells by the reversion of the characteristic hypersensitivity to mitomycin C (MMC), and also by the reduction of the MMC-induced chromosomal aberrations. Gene-editing experiments in WT and FA-A mHSPCs also allowed us to demonstrate TI in these cell types. As it was observed in FA-A MEFs, evidence of phenotypic correction was also observed in FA-A hematopoietic colonies. In these experiments, we also observed a marked toxicity of plasmid DNA nucleofection that compromised the repopulating properties of gene-edited mHSPCs in irradiated recipients.

Altogether, our results demonstrate for the first time the feasibility of implementing a therapeutic targeted integration strategy in a *safe harbor* locus of MEFs and mouse hematopoietic progenitors from a mouse model of Fanconi anemia. Additionally, our data suggest the necessity of limiting the toxicity associated to the nucleofection of mHSPCs with the TALEN and donor as plasmid DNA in order to improve the repopulating potential of gene-edited HSCs.

III. RESUMEN

RESUMEN

La anemia de Fanconi (AF) es una enfermedad hereditaria caracterizada por fallo de médula ósea y predisposición a cáncer. Esta enfermedad está causada por mutaciones en cualquiera de los 20 genes *FANC* descubiertos hasta la actualidad que participan en una vía de reparación del ADN conocida como ruta de AF/BRCA. Las aproximaciones de terapia génica (TG) con células madre hematopoyéticas (CMHs) autólogas podrían constituir un tratamiento eficaz para los pacientes con AF. De hecho, se están desarrollando ensayos de TG convencional no dirigida basados en el uso de vectores lentivirales. Puesto que el riesgo de oncogénesis insercional no puede descartarse por completo en ninguno de los ensayos de TG llevados a cabo en la actualidad con vectores integrativos, se han propuesto nuevas aproximaciones de TG dirigida como alternativas más seguras. Estas estrategias se basan en el uso de nucleasas artificiales que generan roturas de doble cadena en localizaciones específicas del genoma que pueden ser reparadas por la ruta de recombinación homóloga de las células. Esta ruta ha sido utilizada para facilitar la integración, en sitios específicos del genoma, de construcciones donadoras que contienen el gen de interés flanqueado por dos brazos de homología.

En este estudio hemos llevado a cabo una estrategia de edición génica con el objetivo de integrar genes marcadores y terapéuticos en el locus *Mbs85* de ratón, ortólogo al *locus seguro AAVS1* humano. Con el fin de lograr nuestro objetivo, se nucleofectaron las nucleasas TALE (TALEN) junto a diferentes construcciones donadoras que contenían el gen terapéutico humano *FANCA* o el gen marcador *EGFP*, ambos bajo la regulación del promotor de la fosfoglicerato quinasa. Esta estrategia de edición génica se realizó por medio de nucleofección de las TALEN y del donador como ADN plasmídico en fibroblastos embrionarios de ratones con AF-A (*Fanca*$^{-/-}$) y en células madre y progenitores hematopoyéticos (CMPHs) de ratones con AF-A y ratones sanos (WT).

Hemos demostrado por primera vez la actividad de corte de las TALEN diseñadas para cortar el locus *Mbs85* de fibroblastos embrionarios AF-A y de CMPHs de ratones con AF-A o sanos (WT). En los clones editados genéticamente de fibroblastos embrionarios AF-A se demostró integración dirigida, así como expresión de hFANCA. Asimismo, se demostró evidencia de corrección fenotípica en estas células por la reversión de su característica hipersensibilidad a mitomicina C (MMC), y también por la disminución de aberraciones cromosómicas inducidas por MMC. Los experimentos de edición génica en CMPHs de ratones con AF-A o sanos (WT) nos permitieron demostrar integración dirigida en estos tipos celulares. Tal y como fue observado en los fibroblastos embrionarios AF-A, se demostraron también evidencias de corrección fenotípica en colonias hematopoyéticas AF-A. En estos experimentos, también observamos una toxicidad acusada debida a la nucleofección de ADN plasmídico que comprometió las propiedades de repoblación de las CMPHs editadas genéticamente en receptores irradiados.

En conjunto, nuestros resultados demuestran por primera vez la posibilidad de llevar a cabo una estrategia terapéutica de integración dirigida en un *locus seguro* de fibroblastos embrionarios y progenitores hematopoyéticos de un modelo de ratón de anemia de Fanconi. Además, nuestros datos muestran la necesidad de limitar la toxicidad asociada a la

nucleofección de CMPHs con las construcciones TALEN y los donadores en forma de ADN plasmídico para mejorar el potencial de repoblación de las CMHs editadas genéticamente.

IV. INTRODUCTION

1. FANCONI ANEMIA

1.1. General characteristics of Fanconi anemia disease

Fanconi Anemia (FA) is a rare inherited disorder that was first described in 1927 by a Swiss pediatrician named Guido Fanconi [174, 316]. FA is a highly heterogeneous genomic instability disorder and cells from FA are characterized by the accumulation of DNA damage at an increased rate. Mutations in any of the 20 genes up to now described (**Table S1**) have been found as the cause of FA [558]. Depending on the mutated gene, patients are assigned to different complementation groups (FA-A to FA-V). FA is inherited in an autosomal-recessive manner with the exception of the FA-B complementation group, which is X-linked [343]. Fanconi anemia genes work together in an interstrand cross-link (ICL) repair (also known as the FA/BRCA) pathway [106, 118, 278, 363]. Mutations in any of these genes increase the genetic instability of cells, rendering the cells more sensitive to interstrand cross-linking agents. Only 15 out of the 20 genes are considered *bona fide* FA genes (*FANCA, B, C, D1, D2, E, F, G, I, J, L, N, P, Q* and *T*) that satisfy the criteria of the existence of at least two patients who present bone marrow failure (BMF) and whose cells are positive to the chromosomal fragility test [39]. Mutations in *FANCO*, *R* and *S* genes result in a chromosome fragility syndrome without BMF, and are known as FA-like genes. It has been proposed that *FANCM* should be excluded from the list of FA genes, because there is only one FA-M patient that also carries biallelic *FANCA* mutations and the ectopic expression of FANCM in the patient's cells did not rescue the FA phenotype [493]. In addition, the loss-of-function of FANCM described in two Finnish patients did not show any hematological defects [302]. Other FA associated proteins (FAAP) of the FA/BRCA pathway have not yet been shown to be mutated in FA patients: FAAP10, FAAP16, FAAP20, FAAP24, FAAP100 and FAN1 [264].

FA has an incidence of 1-3 per 500.000 live births in the general population, with a carrier frequency of 1/300 [56]. The most frequently FA gene which is mutated in the general population and in the Spanish population is *FANCA* (60% and 80%, respectively). Highly consanguineous groups such as the Spanish gypsies, who share an ancestral mutation in exon 4, present the highest frequency of FA carriers so far described (1 per 67 live births) [56, 64].

1.2. Relevance of the Fanconi/BRCA pathway in genomic stability

The Fanconi anemia/BRCA (FA/BRCA) pathway is activated not only by ICL damage, but also by other DNA damaging agents, such as ultraviolet radiation (UV)[140], ionizing radiation (IR) [42, 290], hydroxyurea [95, 366], and during replication repair of many types of lesions [250]. With the exception of ICLs lesions, all these lesions can also be repaired by other DNA repair pathways [478]. FA/BRCA pathway works coordinating the actions of multiple classical repair pathways in order to remove ICLs, including nucleotide excision repair (NER), translesion synthesis (TLS), and homologous recombination mediated repair (HRR). This fine-tuning regulation also involves the suppression of the non-homologous end joining (NHEJ) pathway [118, 278].

1.2.1. Repair of ICLs by the FA/BRCA pathway

ICL lesions may be formed by endogenous metabolites, such as reactive aldehydes (by alcohol detoxification, histone demethylation, or lipid peroxidation), or exogenous

chemotherapeutic drugs, including mitomycin C (MMC), diepoxybutane (DEB) or cisplatin [73, 174]. ICLs are recognized and excised during all phases of the cell cycle. Nevertheless, FA proteins play a critical role in the replication-dependent repair of ICLs that is required for cell cycle progression into S phase, when the replication fork stalls at an ICL [73, 450]. However, cells that are in G0/G1 phases and that have not yet acquired sister chromatids are forced to repair ICL lesions using the error-prone replication and recombination-independent pathway. In this pathway, ICLs are repaired via unhooking of the one strand via NER followed by TLS to fill the gap, and then a second round of NER to completely remove the lesion.

Specifically, the FA/BRCA proteins that participate in ICL repair have been grouped into three categories: (1) the FA core complex, (2) the ID2 complex, and (3) the downstream proteins that possess a DNA repair function [320] (**Figure 1**).

The FA core complex is formed by three different subcomplexes: (A) FANCL, FANCB, and FAAP100; (B) FANCA, FANCG (XRCC9), and FAAP20 [7, 175, 292]; and (C) FANCC, FANCE, and FANCF [236, 279, 306, 340]. The first protein of the core complex that initiates DNA repair is FANCM (which has ATP-dependent DNA translocase activity), that is stabilized and induced by the MHF complex (MHF1-MHF2-FAAP24)[267, 494, 524]; recognizes the stalled DNA replication fork, recruits the FA core complex to the ICL [230], and generates DNA damage signaling via ataxia telangiectasia and RAD3-related (ATR) kinase [33, 76, 117, 177, 231].

Once the core complex is recruited to the chromatin, FANCL, which contains an E3 ubiquitin ligase, in conjunction with the E2 conjugating enzyme FANCT (UBE2T), monoubiquitinates FANCD2 and FANCI [219, 444]. FANCD2-I monoubiquitination is necessary to the formation of the FANCD2-I complex [178] and localization onto chromatin [243, 319, 475]. While FANCD2 ubiquitination is indispensable for the DNA repair process, FANCI ubiquitination is not essential, although it may enhance repair. Mutations in any of the FA genes of the core complex, except in FANCM [232], prevent the mono-ubiquitination of FANCD2 and FANCI, this being an important event in the regulation of FA/BRCA pathway .

Lastly, the third complex is recruited (from FA-D1, J, N, O, P, Q, R, S and V) together with other DNA repaired molecules. The loss of function in any of these proteins can lead to FA but does not affect FANCD2 ubiquitination[118], so they are known as downstream effectors of the FA pathway [363]. Once in the repair foci, FANCD2/I recruits FANCP that serves as a master scaffold to enhance the activity of a specific complex of nucleases (FANCQ-ERCC1, MUS81-EME1, SLX1 and SNM1A and SNM1B) and incisions in both the 5' and the 3' extremes are generated on the parental strand, allowing the unhooking of the DNA where the ICL has occurred [593]. Unhooking converts the stalled replication fork into a double strand break (DSB) and allows TLS polymerases to repair one strand (Rev1 and Pol ζ- being part FANCV-) [35, 171, 395]. TLS polymerases need the monoubiquitination of the proliferating cell nuclear antigen (PCNA), a polymerase factor that encircles DNA and functions as a moving platform for DNA synthesis, to be recruited to the lesion [362, 540]. The repaired strand can be used as a template using HRR [433]. The activation of the ID2 complex also moves FANCD1 to chromatin, interacting with FANCJ, FANCN, FANCO and FANCS, facilitating the assembly of DNA damage inducing FANCR (RAD51) nuclear foci [281, 476, 501, 515, 573, 576] and the resolution of DSBs by HRR.

Finally, this process generates a fully replicated molecule, but there is still an unhooked ICL that would be removed by NER. Moreover, elimination of the ubiquitin from the ID2 complex is catalyzed by the USP1-UAF1 deubiquitinase (DUB) (that also targets PCNA [394, 499]), this step being critical for the correct function of the FA/BRCA pathway [561].

The main upstream regulators of the FA/BRCA pathway are ataxia telangiectasia mutated (ATM) and ATR-CHK1 cell cycle checkpoint kinases [11, 490, 561]. Specifically, ATR-CHK1 is activated during replication stress and after ICLs lesions [11, 427, 500]. These kinases phosphorylate multiple proteins of the FA pathway, including FANCA/ E/ G/ I/ D2/ S [13, 64, 68], amplifying DNA damage signalling [564]. Additional interactions have been reported for proteins such as bloom-syndrome recq like helicase (BLM), Nijmegen breakage syndrome 1 protein (NBS), and H2A histone family member X (H2AX) [544].

1.2.2. Crosstalk between FA/BRCA and other DNA repair pathways

1.2.2.1. Nucleotide excision repair (NER)

This is a multistep process and versatile mechanism of DNA repair, recognizing lesions produced by a wide spectrum of DNA lesions (UV or IR lesions, ICL agents, reactive oxygen and nitrogen species) [187]. In higher eukaryotic cells, two initial pathways are activated depending on the proteins recognizing the initial damage. The global genome nucleotide excision repair (GG-NER), which detects damage in the entire genome, including untranscribed and silent-chromatin regions; or the transcription-coupled nucleotide excision repair (TC-NER), which repairs transcription-blocking lesions and is involved in replication and recombination-independent ICL repair during G0/G1 phases [145, 563].

The first step is the recognition of the lesion detected by the XPC-hHR23b complex (in GG-NER) or the RNA pol II complex-CS factors (in TC-NER). Then, two transcription factors with helicase activity mediate the strand separation at the side of the lesion. After that, other proteins verify (XPA) and stabilize (replication protein A, RPA) the opened DNA, facilitating the localization of endonucleases (XPF-ERCC1 or XPG) that will remove from 24 to 32 bp of the lesion. Finally, reparative synthesis using the undamaged strand as a template is generated mediated by PCNA processing factor and Pol δ, finishing with a ligation step by ligases (LIG 1/3) [79, 426, 488, 511].

The close interaction between this pathway and FA/BRCA network is observed in diseases in which distortions in NER activity result in UV-sensitive and high-carcinogenic pathologies. Mutations in XPF can provoke three different disorders: xeroderma pigmentosum (XP) [527], progeroid syndrome [393], or FA [38]. Moreover, RPA, which is involved in HRR and FA pathways, connects the initiation of NER to completion of the DNA repair synthesis [407].

1.2.2.2. Translesion synthesis (TLS)

This is a DNA damage tolerance (DDT) mechanism that prevents prolonged replication stalling (in response to UV-induced damage and ICLs lesions) and ensures the completion of DNA replication through the activity of error-prone polymerases, allowing the replication but without correcting the lesion [87]. The low-fidelity polymerases work in a multi-step process.

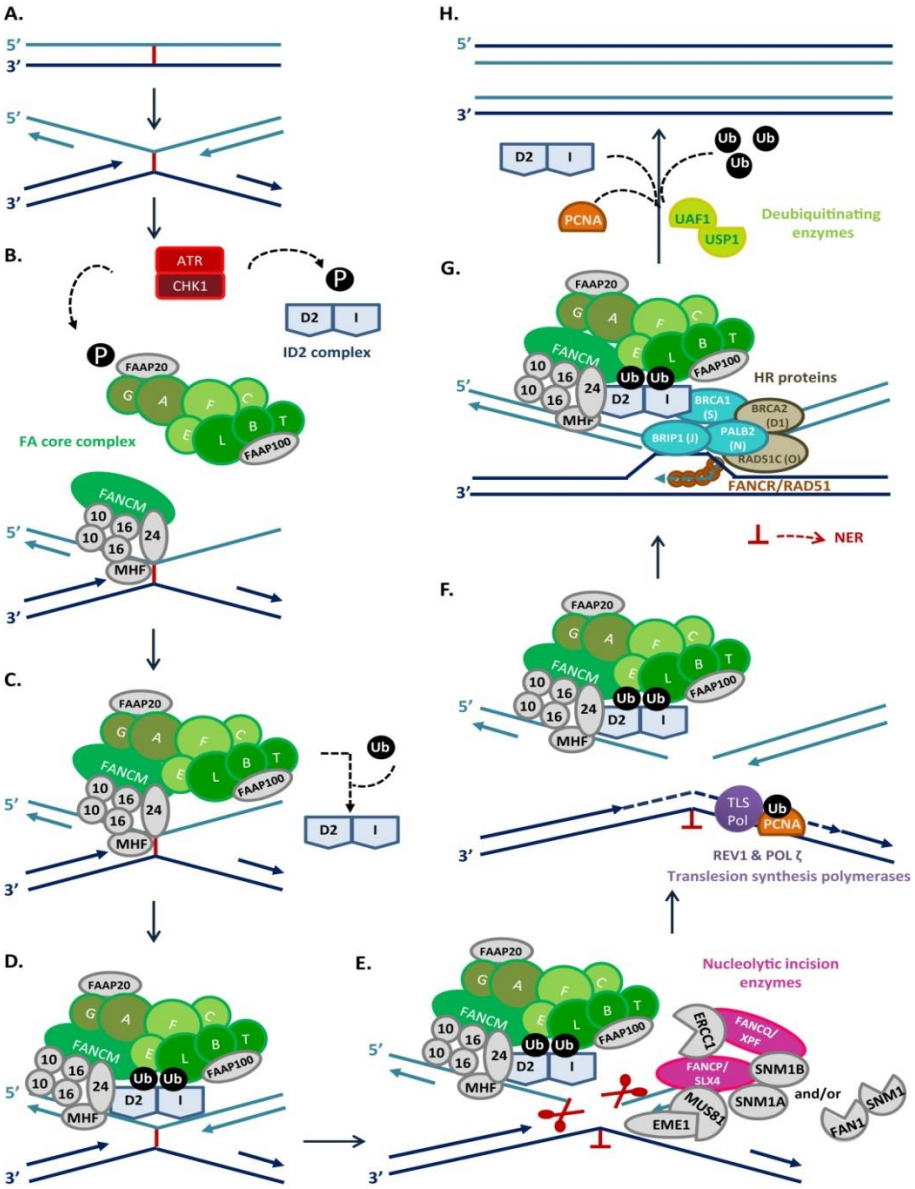


Figure 1. Current model of ICLs repair by the FA/BRCA pathway. **A)** Replication forks that localize on the ICL lesion that covalently links the two strands of the DNA. **B)** FANCM, together with other FA-associated proteins (in grey), activates the DNA damage response mediated by ATR-CHK1 (in red) -that mediates activation by phosphorylation of some proteins of the pathway- and facilitates the recruitment of the rest of the FA core (in green) to the lesion. **C)** Monoubiquitination of the FANCD2/I complex (in blue) mediated by FANCL/T. **D)** FANCD2/I-Ub complex moves to the DNA lesion. **E)** FANCD2/I-Ub act as a platform to recruit multiple nucleases to coordinate nucleolytic incisions flanking the ICL: FANCQ, FANCP (in pink), and other proteins (in grey) to unhook the ICL. **F)** Translesion polymerases (in purple) repair the leading strand (assisted by PCNA-Ub). **G)** The repaired strand serves as a DNA template to repair the other strand by homologous recombination (HR) proteins (in brown and blue, RAD51 as a brown circles) **H)** USP1-UAF1 (in light green) deubiquitinate the ID2 complex (and PCNA), having resolved the ICL. Modified from Kim et al. [264].

Following replication arrest, E2-E3 ubiquitin ligase complex (RAD6-RAD18) monoubiquitinates PCNA, recruiting insertion TLS polymerases (from Y- family, notably Rev1) to the lesion that incorporate a nucleotide opposite to the damage site. Then, an extension polymerase (Pol ζ complex) extends the TLS patch by 18 bp. Finally, Pol ζ is changed by a high fidelity DNA polymerase that processes DNA replication [206, 264].

The TLS and FA/BRCA pathways are highly coordinated. RAD6-RAD18 can activate the FA network through FANCD2 ubiquitination [165, 183, 592]. In addition to this, PCNA acts as a scaffold protein that by promoting FANCD2 ubiquitination, recruits both FANCL and FANCD2 to the chromatin [183]. Furthermore, once FANCD2 is ubiquitinated, it recruits Pol η to the lesion, facilitating TLS polymerase switch, and RAD18 protein is recruited to the single-strand DNA (ssDNA) in a RPA dependent manner, a process that links TLS and HRR [110]. FAAP20 binds with Rev1 by a ubiquitin-binding zinc finger (UBZ) domain [266]. Finally, PCNA is also deubiquitinated by USP1, as happens with the ID2 complex, demonstrating that the regulation of both pathways should be very similar [234].

1.2.2.3. FA/BRCA and Homologous Recombination mediated Repair (HRR) and Non-Homologous End Joining (NHEJ)

These two mechanisms involved in the repair of DSB are explained in detail in Introduction section 3.1. The coordination between the FA network and HRR is evident from the fact that many FA proteins are also involved in the HRR step during ICL repair, and their deletion renders cells hypersensitive to ICL-inducing agents. The absence of a functional FA pathway is speculated to be the cause of the gross chromosomal abnormalities observed in FA cells due to the activation of the error-prone NHEJ repair. Moreover, inhibition of the NHEJ pathway in cell lines derived from patients with FA can reduce the toxicity of ICL-inducing drugs, indicating a role of FA pathway in suppressing NHEJ [1, 408]. FANCM also plays a role in HRR repair independent of the FA network, as its depletion reduces RPA foci formation in the 3'-ssDNA overhang [231]. In addition, FANCG interacts with BRCA2, leading to the colocalization of the core complex with RAD51 in the DNA lesion, and inactivation of FANCG decreases HRR [238, 581]. Furthermore, BRCA1 inhibits NHEJ dependent on DNA-PKcs, also excluding P53BP1 accumulation [1, 9, 52, 408]. FANCD2 is needed for proper recruitment of the CtIP protein (involved in DNA end-resection during HRR) at the ICL [142, 375, 541] and its monoubiquitination increases the HRR 2-fold [381].

1.3. Phenotypic manifestations in Fanconi anemia cells

FA cells present a characteristic phenotype that may be produced by the combination of different factors including a defect in the repair of ICLs lesions that leads to the accumulation of DNA damage and genomic instability, but also due to increased levels of oxidative stress [392, 410, 415]. The main features of FA cells are summarized in **Table 1**.

1.4. Clinical aspects and diagnosis of Fanconi anemia

In spite of the high heterogeneity of the disease, most FA patients show birth defects and BMF, and have a predisposition to cancer. 70% of patients present congenital abnormalities including skeletal abnormalities; radial axis defects; skin pigmentation anomalies; short stature; small head and eyes; problems in heart, kidney, intestine and urinary tract; small reproductive organs and infertility in males and reduced fertility in females; mental retardation or learning disabilities; and low birth weight [18, 482, 533].

The main hallmark of the disease is BMF [70, 258] that generally appears between in the first 10 years of life, and takes place in~ 80% of 15-year-old patients, being the main cause of death. Macrocytosis is usually the first symptom, followed by thrombocytopenia and neutropenia [55, 280].

Feature	Description
1.-Hypersensitivity to ICLs agents such as MMC or DEB	Exposure to these agents induces chromosomal aberrations. This is used as a hallmark for FA diagnosis in peripheral blood (PB) T lymphocytes [16, 18, 411].
2.-Defects in monoubiquitination of FANCD2 and recruitment to the DNA repair foci	Cells with mutations in the core complex proteins are not able to ubiquitinate D2, and this is used as a diagnostic marker [428, 503]. Monoubiquitinated FANCD2 could be observed by Western Blot (WB) or by foci formation [176].
3.-Cell cycle disturbance	Cells induce a G2/M cell cycle arrest after ICLs exposure due to the high number of stalled replication forks. The G2/M arrest in fibroblast treated with ICL agents is useful for diagnosis of mosaic patients [392]. If the G2/M cell cycle arrest is overcome, the strong induction of p53 and p21 gives rise to a new cell cycle arrest in G0/G1. This mechanism has been proposed as the origin of the loss of hematopoietic stem and progenitor cells (HSPCs) leading to bone marrow failure (BMF) [68, 139].
4.-Oxygen sensitivity and increased levels of ROS	This results from the deregulation of reactive oxygen species (ROS) metabolism. It produces defects in the culture of cells, their culture in hypoxic conditions being necessary [248, 409, 430, 480]. Moreover, some FA proteins (A/ C/ G/ D2/ J) participate in redox and detoxification metabolism [132, 133, 411, 487, 543, 595]. Products of endogen metabolism, such as acetaldehydes, could induce DNA mutations, inducing BMF and cancer predisposition [91, 173, 287, 288, 461].
4.-Cytokine sensitivity	FA cells produce high levels of pro-apoptotic and pro-inflammatory cytokines such as interferon γ (INF-γ) [137], interleukin 6 (IL-6) [146], tumor necrosis factor α (TNF-α)[134, 464, 484] or interleukin 1β (IL-1β) [239]; their overexpression has been associated with a pro-apoptotic phenotype involved in BMF, developmental malformations and leukemogenesis [146, 296, 338].
5.-Telomere abnormalities	FA cells have shorter telomeres than those of healthy individuals, but the telomerase activity is probably increased to compensate for this modification [55, 249, 291].
6.-Defective adhesion and homing ability	In FA cells, CDC42 ATPase is downregulated by over-expression of TNF-α, and thus cell migration is affected [596].
7.-Defects in cytokinesis	The number of anaphase ultra-fine bridges is increased in FA cells during replication stress, increasing the risk of cytokinesis failure. Therefore, the number of binucleated cells and supernumerary centrosomes are also multiplied [84, 336, 385, 552].

Table 1. Characteristic phenotype of FA cells.

Furthermore, FA patients present a cumulative incidence of hematologic malignances, presenting acute myeloid leukemia (AML) or myelodysplastic syndrome (MDS) with a risk of 30-55% in 40 year-old patients. In addition to this, 29-30% of the 40-year-old FA patients are prone to develop specific solid tumors, such as head and neck squamous cell carcinomas (SCC), esophagus SCCs, liver tumors and vulvar cancer in women [9-11, 75, 456].

Due to the genetic instability of FA cells, some patients can undergo "somatic mosaicism" when the spontaneous reversion of the pathogenic mutation occurs in a somatic cell, such as in T cells. These mutations could lead either to the wild type (WT) form of the FA gene, or to compensatory mutations, that generate a functional FA protein. If the reversion takes place in a progenitor cell or in a hematopoietic stem cell (HSC), the proliferation advantage of these cells could lead to the partial or total recovery of PB cell counts in FA patients [191, 194, 315, 331, 503, 555]. The diagnosis of mosaic FA patients (which comprise up to 25% of FA patients) is often difficult because in some instances they do not show clear signs of the disease [18, 315].

In addition to this, at the time of diagnosis it is essential to discriminate FA from other diseases with similar phenotypes [126, 516]. The differential test for the diagnosis of FA patients is the chromosomal breakage analysis induced with ICL agents in PB T lymphocytes (such as DEB or MMC) [16, 483]. A complementary test of G2/M cell cycle arrest induced by ICL agents is useful to confirm the FA diagnosis of mosaicism in fibroblasts [392]. These studies should also be

performed in apparently healthy FA patient siblings because they could be undiagnosed. Next, FA complementation groups screening is performed to detect the mutated gene and to assign the patient to a specific complementation group. Typically, PB T cells (or fibroblasts from mosaic patients) are transduced with a battery of gammaretroviral (γ-RVs) or lentiviral (LV) vectors encoding different FA genes and then treated with ICL agents. The vector that reverts the sensitivity determines the FA complementation group [64]. As long as none of the most common FA genes (*A, C, G, E, F,* and *D2*) complements the mutation, analysis by WB of the monoubiquitination of FANCD2 allows scientists to define if the FA affected protein is upstream or downstream of the ID-complex. After that, RV subtyping of the non-common FA upstream ID-genes or molecular and functional analyses of downstream proteins are performed. New approaches such as whole genome sequencing have proved to be very efficient in the testing process [88], allowing the identification of new genes, as has been recently shown [38, 557]. In addition, further sequencing, RNA analysis, and SNP chips could be helpful to identify a second mutation [86, 597].

1.5. Fanconi anemia mouse models

1.5.1. Description of the FA mouse models

Laboratory mouse models represent a powerful tool to understand the FA clinical phenotypes and to evaluate the efficacy and the toxicity of potential treatments. The current FA mouse models do not fully recapitulate the main features of the disease, but most of them show reduced fertility [420, 532], and cells derived from most of all mutant mice are sensitive to ICL agents. The viability and phenotype of the disease in these models could be affected not only due to the function of the disrupted gene, but also to their genetic background (especially the C57BL/6 strain is more lethal). Moreover, it has been proposed that the controlled environmental conditions of the laboratory prevent the disease from evolving; only under harmful genotoxic stresses (such as exposure to MMC or ethanol, or to endogenous molecules, such as acetaldehyde or ROS) do FA mice acquire some of the pathogenic characteristics of the FA disease [23, 570]. **Table S2** summarizes the current FA mouse models and their characteristics.

Briefly, embryonic and perinatal lethality have been observed in models with disruption in genes such as *Fancd1*$^{\Delta 27/\Delta 27}$, *Fancd2*$^{-/-}$, *Fancl*$^{-/-}$, *Fanc*m$^{-/-}$, *Fanc*n$^{-/-}$, *Fanco*$^{-/-}$, *Fancp*$^{-/-}$ and Usp1$^{-/-}$ (**Table S2**). The loss of function of *FANCD1* and *FANCN* is also lethal in humans, and therefore some of these mouse models are hypomorphic rather than knock-out [150, 531]. In other models, embryonic lethality has been circumvented with conditional models [251]. Furthermore, most FA mouse models do not display any specific developmental abnormality. Some of them develop microphthalmia, and *Fancp*$^{-/-}$ mice have abnormally-shaped and enlarged skulls [60, 83, 223, 268, 571]. Of the FA mouse models harbouring a unique mutated gene, none reproduces the BMF observed in patients, although some of them present leukopenia and thrombocytopenia, proliferation defects in hematopoietic progenitors and derived hematopoietic colonies, or reduced repopulating ability (**Table S2**). Specifically, *Fancp*$^{-/-}$ model [83] and *Fancd1*$^{\Delta 27/\Delta 27}$ hypomorphic model [339, 386] are the models that most closely simulate the hematologic phenotype of FA patients. Moreover, although some FA models have been reported to present a significant incidence of tumors, it has been described that it is very rare that FA mouse models spontaneously develop cancer. Aldehydes and other genotoxic stresses could contribute to the pathology of FA. Double mutants for *Fancd2* and *Aldh2* and for Superoxide

dismutase 1 (*Sod1*), as well as p53/FA double mutants present more severe phenotypes [173, 203], these last ones being models of tumorigenesis (**Table S2**).

1.5.2. *Fanca*$^{-/-}$ (FA-A) mouse model

This model was created by the deletion of exons 4-7 by Cheng et al. [97] and, contrary to the case of most FA-A patients, it presents a mild phenotype. Mouse embryonic fibroblasts (MEFs) derived from these mice presented MMC hypersensitivity; and both male and female mice showed reduced fertility due to hypogonadism. Moreover, these mice did not spontaneously develop congenital abnormalities. In our laboratory, Rio et al. [455] reported that *Fanca*$^{-/-}$ mice develop thrombocytopenia in a C57BL/6 background. In addition, megakaryocyte progenitors, but not granulocyte-macrophage progenitors from the bone marrow (BM) of *Fanca*$^{-/-}$ mice, had impaired proliferation *in vitro,* and were highly sensitive to MMC. Finally, it was demonstrated that this phenotype could be rescued by transduction of the hematopoietic progenitors with a vector expressing the human *FANCA* cDNA [365, 455]. In this work, we have decided to work with this model because the *FANCA* gene is the most frequently mutated worldwide in FA patients, and because FA-A cells are compatible with HR-mediated gene-editing approaches [456].

1.6. Treatments for Fanconi anemia patients

1.6.1. Conventional treatments for Fanconi anemia patients

FA patients are normally followed by measuring their blood cell counts and by analysing their BM to monitor the progression of BMF, AML and MDS. BMF is the main cause of mortality in FA patients, and hematopoietic stem cell transplantation (HSCT) constitutes the preferred curative therapy for these syndromes. However, other palliative treatments are available to provide temporary hematopoietic support when residual endogenous hematopoiesis remains.

The possible palliative treatments are based on the use of hematopoietic growth factors, androgens, and transfusions [138] to sustain acceptable numbers of cells in PB. Growth factors are no longer used because of their short term response and due to concerns about their contribution to hematopoietic cell transformation [196, 438, 463, 530]. However, the majority of patients do respond to androgen therapy, which is useful in early stages of BMF. Oxymetholone has been the most frequently used, although due to some side effects associated with it [549], danazol is nowadays more frequently used in FA patients [479].

Currently, HSCT is the only curative treatment for the hematologic signs of FA, either using BM or mobilized peripheral blood. HSCT offers high probabilities of cure when human leukocyte antigen (HLA)-identical sibling donors are available [323, 423, 517]. However, the chances of having HLA-identical donors are reduced. Moreover, HSCT has been challenging due to the poor tolerance of FA patients to conditioning regimens involving alkylating agents and to the high incidence of graft-versus-host disease (GvHD). New reduced conditioning regimens, the addition of immunosuppressive drugs such as fludarabine, or alloreactive T-cell depletion strategies have reduced the incidence of GvHD in alternative HLA-matched related or unrelated transplants [40, 195, 324, 523, 587]. Furthermore, recent advances in prenatal HLA-typing and *in vitro* fertilization and preimplantation genetic diagnosis (IVF-PGD) followed by embryo

selection offer new alternatives to identify a matched sibling donor whose HSCs could be the source of transplant for the FA sibling patient [31, 192, 252, 538, 550].

The non-hematologic complications of FA patients should also be addressed. FA patients have *per se* an increased tendency to develop solid tumors which is incremented after HSCT. To prevent leukemias, BM aspirations should be performed periodically to rule out premalignant clonal expansion, and head and neck evaluation, as well as gynecologic exams are required to prevent SCCs.

1.6.2. Innovative therapies for Fanconi anemia patients

Due to the difficulty of finding HLA-matched donors, and to the risks associated to allogenic HSCTs, gene therapy approaches with autologous cells might constitute a potential therapy for FA patients. In the next sections, new therapeutic options for FA patients, involving conventional (also known as non-targeted) gene therapy and/or targeted/directed gene therapy approaches will be explained.

2. GENE THERAPY

Gene therapy (GT) is a novel therapeutic option consisting of the delivery of genetic material into specific cells affected by a disease [257]. The correction could be carried out by different approaches (**Figure 2**). If the integration of the therapeutic gene in the target cell occurs without eliminating the defective gene, the GT approach is known as a **gene addition**. This could be performed by using the HRR machinery to introduce the therapeutic gene into a *safe harbour* locus, or by semi-random integration. When the mutant copy is substituted specifically by the therapeutic gene, the approach is known as **gene replacement** either by means of **gene correction** or by **knock-in,** which involves HRR to replace either small mutations or the mutated cDNA. Additionally, a **suppression strategy** can be used to eliminate (knocking-out, involving gene targeting) or to reduce (knocking-down, using interference ribonucleic acid, RNAi) the expression of a malfunctioning gene. This strategy could also be applied to generate models of disease by disrupting specific genes.

If the integration of the genetic material is not directed to a specific locus, it is referred to as **conventional** or **non-targeted GT,** whereas if it is directed to a specific locus it is called **targeted GT** and requires the use of gene-editing tools. Moreover, depending on the modality of manipulation and on the treatment of the cells, the introduction of the therapeutic gene could be performed by ***in vivo*** administration of the vehicle that contains the therapeutic transgene into the patient, or by ***ex vivo*** manipulation of the cells with the therapeutic vehicle, followed by their re-infusion into the patient.

2.1. Relevant considerations for the application of GT protocols

Several important aspects should be taken into account to achieve an efficient and safe GT protocol.

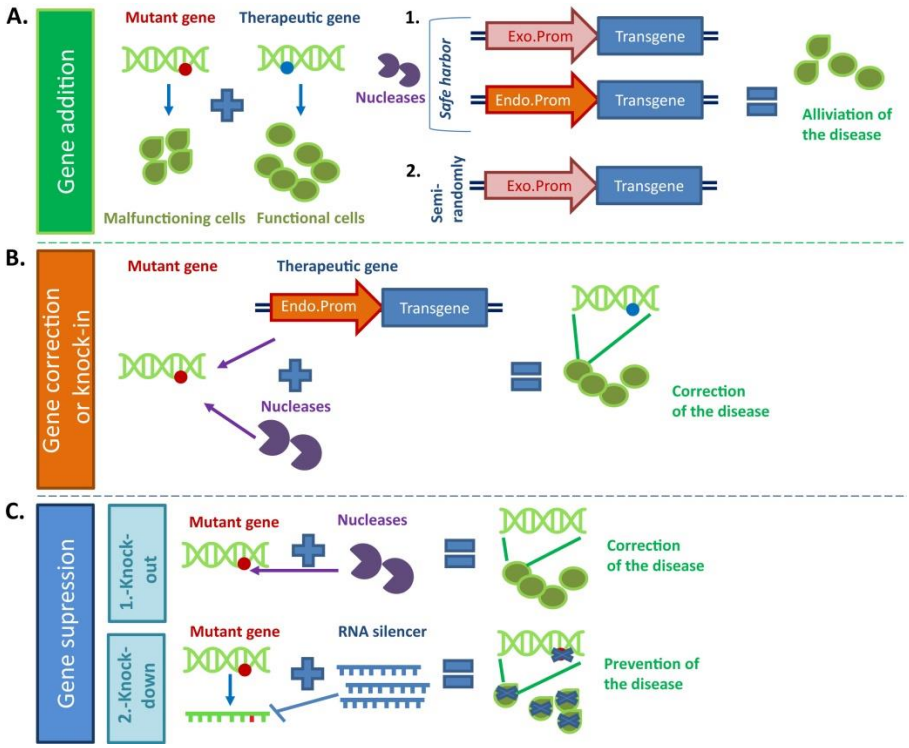


Figure 2. Different GT therapy strategies. A gene mutation (red dot) abrogates protein synthesis and leads to disease (abnormal shaped cells). **A) Gene addition strategy** corrects the disease: **(1)** By HRR using gene-editing tools such as nucleases (purple pie) to directly cleave into a *safe harbor* locus and by providing a functional copy of the therapeutic cassette whose expression could be regulated by an exogenous or endogenous promoter. **(2)** By integration of the therapeutic cassette in a semi-random fashion. Both strategies alleviate the disease by supplementing the therapeutic gene (blue dot in a different location) that ameliorates a specific aspect of the disease, giving rise to functional cells. **B) Gene replacement strategy** that corrects the disease by HRR substituting just some of the mutated nucleotides or the mutated cDNA giving rise to functional cells. **C) Gene suppression strategy** corrects the disease: **(1)** By knocking-out, using gene-editing tools disrupting or eliminating a disease-promoting sequence to a disease-free sequence. **(2)** By knocking-down, using interference RNAs, RNAi (blue combs) to inhibit the aberrant messenger ribonucleic acid, mRNA (green comb), thus preventing the synthesis of toxic protein aggregates and correcting the disease. Modified from Wang et al. [559].

a) Selection of the transfer vector

The transfer of exogenous genetic material can be performed by different methods. In most cases it implies the use of a vehicle to facilitate the introduction of the genetic material. The vehicles used in GT can be classified into two large groups: a) viral vectors (based on genetically modified viruses) and b) non-viral-based systems.

Traditionally, viral vectors based on γ-RV or LV [258, 326] have been used for *ex vivo* GT of HSCs. Other viral vectors are applied however, in other fields of GT, based on adenovirus (AdV) or adeno-associated virus (AAV) [82, 108]. Gene transfer methods based on non-viral vectors [190, 260] such as naked DNA [308, 569], minicircle DNAs [96, 181, 259], DNA or mRNAs complexed to different macromolecules [44, 391] or to metal nanoparticles [107, 124], and transposons [123, 245, 572] are also used. These technologies could be delivered by various routes including intravenous infusion, oral ingestion, hydrodynamic transfection [214, 307], or electroporation [209, 286]. Viral systems offer high efficiencies and long-term expression, the possibility of infecting dividing and non-dividing cells (depending on the viral vector) and can integrate or remain in episomal state, whereas non-viral systems are very simple, could transfer large molecules and are less immunogenic.

b) Target cells

HSCs have been proved to be ideal target cells, not only due to their potential to self-renew and to give rise to all hematopoietic lineages, but also due to the possibility of isolating and genetically correcting them *ex vivo,* and re-infusing them through intravenous infusion [185]. Furthermore, thanks to the recent improvements in HSCT-associated technologies, *ex vivo* GT protocols have been developed with autologous HSCs, and these modified cells represent a therapeutic option for several monogenic diseases in which the corrected HSCs could restore PB counts [383]. In the field of immunotherapy, T cells have been modified by GT and nowadays several clinical trials are underway to treat B cell malignancies with anti-CD19 CAR autologous T cells [25, 186].

c) *Ex vivo* culture

The genetic modification of HSCs requires the *ex vivo* manipulation of the cells, which may affect their self-renewal ability, viability, homing and *in vivo* repopulating ability. Thus, short culture times are preferred because they preserve the HSCs characteristics [371].

d) Appropriate and efficient gene expression and selective advantage

GT protocols must provide appropriate levels of transgene expression. For this reason, the transfection or transduction efficiencies and vector choice are critical factors that must be considered. Furthermore, the integration site, local epigenetic features, gene deserts or heterochromatic regions may influence the level of transgene expression. In addition, the choice of the promoter is also crucial in ensuring physiological levels of expression and avoiding the trans-activation of upstream or downstream genes [600].

In certain disorders, cells corrected by GT can develop a proliferative advantage, by which a reduced number of corrected cells would be able to proliferate and restore the affected system. This phenomenon has already been observed in X-linked severe combined immunodeficiency (SCID-X1) [200], in Adenosine deaminase SCID (ADA-SCID) [4, 5, 155] and in Wiskott-Aldrich syndrome (WAS) [141, 554] treated patients. In FA patients, due to the observed mosaicism in some patients and the preclinical experiments already done, it is expected that a similar phenomenon will also occur in the FA patients treated in GT clinical trials [387].

e) Host immune response

One of the biggest barriers of a successful GT is caused by the activation of the host immune system against the transgene and/or vector particles. Designed vectors should circumvent immunological recognition that impairs engraftment. Different strategies, such as alterations of the vector structure [12, 28, 335], the induction of antigen-specific tolerance by *ex vivo* or *in vivo* induced T regulatory cells [460], or the inhibition of the transgene expression in antigen presenting cells [50] have been proposed.

f) Genotoxicity

Insertional oncogenesis is caused by enhancer activation of neighboring genes, mainly proto-oncogenes. This phenomenon was observed during the first GT trials conducted with vectors based on γ-RV that produced hematopoietic malignancies due to the clonal expansion of HSCs containing an integration of the proviral retrovirus. This integration induced the upregulation of proto-oncogenes that led to T-cell acute lymphoblastic leukemia (T-ALL) and

AML in SCID-X1 and WAS patients [47, 198, 225, 274]; or to the upregulation of myeloproliferative genes that led to MDS, such as in chronic granulomatous disease (CGD) patients [193, 358, 368]. Moreover, other factors have been proposed to be involved in the development of hematopoietic malignancies, such as the potency of the promoter/enhancer elements of the vector, aberrant splicing from the vector transcripts [70, 112, 361], synergy with the transgene [274], or the HSCs predisposition to insertional mutagenesis [210].

In order to reduce the risks associated to γ-RV, the inactivation of the U3 region in the 3' end of the viral LTR (long terminal repeats), leading to self-inactivating (SIN) vectors [201] has been applied. Additionally, the use of more physiological or lineage-specific promoters [10], the choice of vectors less prone to integrate near promoters or active genes such as LVs [351], and the incorporation of insulators that protect surrounding genes [384] have been considered.

2.2. Relevant *ex vivo* non-targeted GT protocols with HSCs

Non-targeted *ex vivo* HST GT has become a feasible therapeutic approach for several diseases, particularly for monogenic hematological disorders. Thus, different clinical trials have been conducted that are summarized in **Table S3** [41, 71, 156, 185, 283, 384, 432, 536].

2.3. Non-targeted GT in Fanconi anemia

One of the crucial features of FA that distinguishes it from other monogenic diseases is the marked proliferation defect of their primitive HSCs [17, 68, 289]. Therefore, a survival advantage would be expected from corrected HSCs, potentially allowing normalization of hematopoiesis even in the absence of or after mild conditioning. This fact is supported by the phenomenon of somatic mosaicism observed in some FA patients [191, 194, 315, 331, 503, 555].

Three different clinical trials have been conducted with $CD34^+$ cells transduced with γ-RV for 3 to 5 days, followed by their infusion into non-conditioned FA patients, but in none of these trials were clinical benefits observed [261, 311, 556]. In the published studies [261, 311], the authors transduced BM or granulocyte colony-stimulating factor (G-CSF) mobilized $CD34^+$ cells with FANCC or FANCA γ-RVs [261]. The common factor for both studies resided in the low number of corrected cells that were infused per patient, and whereas in one study a very small number of corrected cells were detected in PB or BM during a certain time, in the other the presence of the therapeutic transgene was not detected, but transient increases in hemoglobin (Hb) and platelets (PLT) counts were observed.

The lack of success in these previous GT trials motivated the creation of an International Fanconi Anemia Gene Therapy Working Group to define the optimal strategy to implement GT clinical trials for individuals with FA [534, 535]. This consortium discussed the choice of viral vector, prestimulation and transduction methods, clinical GMP (good manufacturing practice) vector manufacture, target cell populations and appropriate criteria for the treatment of FA patients [348, 387]. Of particular relevance is the $CD34^+$cells collection method to be used in these patients. In this regard, a phase II clinical trial was opened in 2012 in Spain (Eudra CT number: 2011-006197-88), in which our group is participating, using G-CSF and plerixafor to obtain high numbers of mobilized PB $CD34^+$ from FA patients [125, 222, 436, 583]. Furthermore, preclinical studies of FA GT with LV [27, 170, 189, 247, 374] promoted the set-up of two clinical trials for FA-A patients [13]. In 2012, a phase I clinical trial was opened in the USA (NCT01331018), and up to now two

patients have been recruited, although the results have not been published yet. Moreover, in 2015, a phase I/II GT clinical trial (FANCOLEN: Eudra CT number: 2011-006100-12) was opened in Spain, using another SIN-LV [189, 365] that has been designated as an orphan drug by the European Commission (EMA/COMP/662962/2010) and by the FDA (#16-5193) . In this clinical trial, two patients have been treated so far.

3. TARGETED GENE THERAPY

Although SIN-LVs have clearly improved the safety of hematopoietic GT, some concerns can not be ruled out completely due to their preferential integration into transcriptionally active genes [26]. Targeted GT has evolved to address this fundamental limitation of non-targeted GT. It relies on the use of artificial nucleases that recognize and cleave specific DNA loci, causing DSBs to facilitate the generation of targeted modifications in the genome [72]. These gene-editing modifications take advantage of two major DNA repair pathways of the cell that are activated after DSBs damage: HRR and NHEJ [521].

NHEJ involves the direct religation of the DNA ends without any template, in an error-prone manner, generating insertions and deletions (INDELS) at the break site [300]. NHEJ has been exploited to disrupt or excise abnormal sequences, or to change the reading frame of a gene in a therapeutic context [14, 406, 473, 526]. In contrast, natural HRR repair depends on the strand invasion of the broken end into a homologous sequence and subsequent repair of the break [519]. The low frequency of HRR in mammalian cells (approximately 10^{-6}) [441] has impeded the evolution of this pathway for therapeutic aims [59], but the trend has changed thanks to the studies and improvements made over the last two decades [102, 466, 467, 496] and particularly, due to the development of artificial nucleases. The homology-directed repair (HDR) strategy has been adopted for different aims that are explained in Introduction section 3.2.

3.1. Cellular mechanisms involved in the repair of DSBs

DSBs are considered to be the most deleterious DNA lesions because if they are not repaired they can lead to potential loss of genetic information. Unrepaired lesions could result in cell death, or chromosomal aberrations that may lead to cancer [2, 263, 575]. DSBs could be induced by endogenous sources such as ROS, cellular metabolism and replication-associated errors; or by exogenous sources including IR or chemotherapeutic agents [182, 208]. Typically, two main pathways have been described, HRR and canonical or, DNA-PKcs (DNA-protein kinase) dependent NHEJ (D-NHEJ). In addition, other two error-prone mechanisms, namely alternative or backup NHEJ (B-NHEJ) and single-strand annealing (SSA), have been shown to operate (**Figure 3)**.

3.1.1. Non-Homologous End Joining (NHEJ)

NHEJ is the preferred mechanism of the cells to repair DSBs (up to one thousand fold more frequent than HRR [153]). NHEJ operates with poor fidelity generating INDELs [111]. Two different pathways have been identified: the D-NHEJ and the B-NHEJ pathways.

D-NHEJ presents fast kinetics with the aim of protecting genome integrity and it predominates in G0/G1 phases of cell cycle. In D-NHEJ, the Ku heterodimer binds the two blunt DNA ends, and it recruits and activates the catalytic subunit DNA-PKc, which stabilizes the DNA ends [341]. This complex facilitates the rejoining mediated by a DNA ligase 4/X-RCC4 heterodimer in concert with nucleolytic enzymes (that generate ligatable ends) [111, 567].

The backup pathway was discovered in D-NHEJ-deficient cells [122] and is slower than D-NHEJ. This pathway is potentiated during S and G2 phases, operating as a backup to HRR in these phases. This pathway tends to generate more mutations than the D-NHEJ, frequently generating chromosomal translocations, telomere fusions and other genomic rearrangements such as the switch recombination in cells of the immune system. This pathway usually depends on the presence of microhomologies near the DSB ends, most probably found when DNA ends become resected. The main factors involved are PARP-1, the MRN complex, and CtIP. Other proteins such as WRN form a complex with the ligase LIG3 (or LIG1) [119, 354].

3.1.2. Homologous Recombination mediated Repair (HRR)

HRR is a highly conserved, error-free repair mechanism. This mechanism uses the sister chromatid as a homologous template, hence only being active in the S and G2 phases of the cell-cycle [399, 470]. However, HRR could also be achieved in very rare cases in diploid cells during the G1 phase of cell cycle, using the homologous chromosome as template [81].

HRR requires three major steps: end resection, strand invasion and finally resolution. First, the DSBs are recognized by the MRN, which along with CtIP promote the initiation of the resection. After that, different exonucleases (Exo1, DNA2 and BLM helicase) are recruited by the BRCA1 complex to generate 3'-ssDNA overhang that is coated by RPA protein. Then, RPA is replaced by RAD51 recombinase (requiring other mediators such as BRCA2, RAD51 paralogs and PALB2) that generates a nucleoprotein filament and invades the intact double-strand DNA from the sister chromatid to search for homology and form a transitory structure known as displacement loop (D-loop). When homology is found, RAD51 is released from the 3' overhang of the invading strand, and this overhang is used for primer elongation [299]. HRR can go in several directions from this point. The most frequent in higher eukaryotes is known as synthesis-dependent strand annealing (SDSA), in which the elongation of the invading 3'-end can continue up to a limited distance, followed by displacement of the newly synthesized stretch, and re-ligation with the original DNA resulting in the DSB repair. Alternatively, a second mechanism known as double strand break repair (DSBR) could act when a second end-capture occurs, leading to the formation of a double Holliday junction (dHJ). Resolution of the dHJ is carried out by a group of resolvases, and crossover or non-crossover products can emerge depending upon which strands are digested [69, 355].

3.1.3. Single-Strand Annealing (SSA)

SSA is initiated when DSBs occur at a genomic sequence, where extensive homology exists between sequences at either side of the DSB and mediates rejoining of the two DNA ends. Despite using homologies, SSA is a non-accurate mutagenic process, causing large DNA deletions between the two homologous segments, including the loss of one of these segments [255]. Analogous to HRR, SSA is initiated by a similar model of end resection, but the 3'-overhangs are not initiating repair synthesis with strand invasion [157, 512]. The two homologous stretches

are annealed by RAD52, the generated tails are removed by ERCC1/XPF nuclease complex, and the resulting nicks are ligated by DNA ligase LIG1 [518].

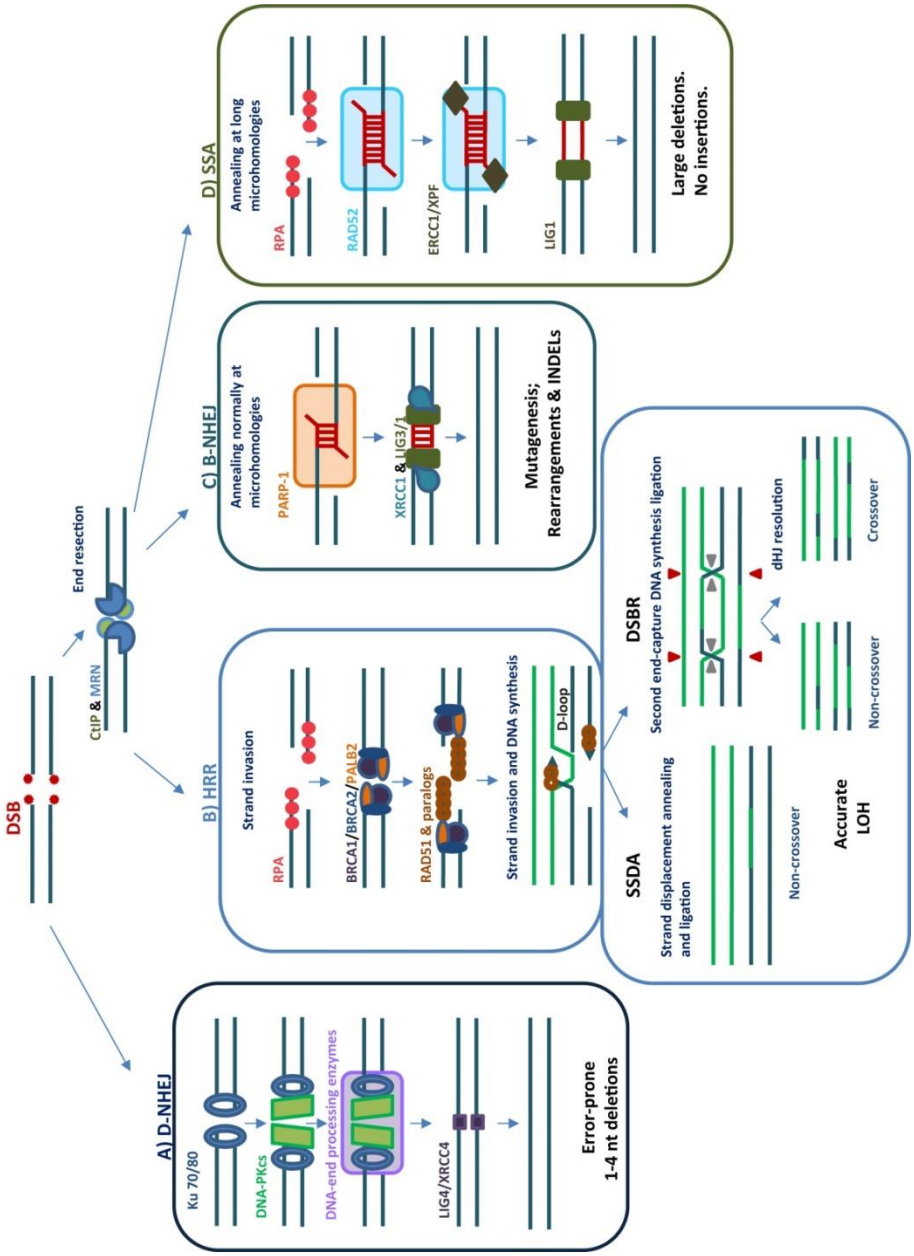


Figure 3. Double-Strand Breaks (DSBs) DNA repair pathways. Some DNA repair pathways are end resection-independent such as D-NHEJ **(A)**, while others are end resection-dependent such as HRR **(B)**, B-NHEJ **(C)**, and SSA **(D)**. After strand invasion, the HRR pathway could go in two different directions: the SDSA, or the DSBR. Each pathway leads to different genetic outcomes with different fidelity. The different proteins involved in each pathway are represented and explained in detail in the text. nt: nucleotides; LOH: loss of heterozygosity; D-loop: displacement loop; dHJ: double Holliday junction. Modified from Ceccaldi et al. [69].

3.1.4. DSB repair pathway choice

The choice of the DSB repair pathway would have the ultimate aim of preserving the genomic integrity after the lesion. Several factors could strongly modulate the pathway: the type of DSB, the presence of non-DSB lesions in close proximity to the DSB, chronic DSBs, the chromatin context, the cell type, the growth-state phase or the presence of growth factors in

the media [75, 355]. But cell cycle and proteins activated during the different phases are clear determinants of the choice. HRR is normally restricted to S/G2 phases due to the presence of the sister chromatid. Moreover, important key factors such as RAD51, BRCA1, BRCA2, BLM and CtIP are only upregulated during these phases [143, 148, 158, 237, 562, 586]. The central step of end-resection is regulated by the interplay between BRCA1 and P53BP1 [43, 45, 52], while BRCA1 supports resection in S/G2 phases, P53BP1 inhibits resection in G1 phase [149]. In contrast to HRR, D-NHEJ operates throughout the cell cycle, although it is preferred in G1 phase [465]. B-NHEJ is activated in cells in G2 phase, deficient in D-NHEJ that fail to repair DSBs by HRR [241]. SSA is strongly suppressed by HRR and vice-versa, inactivation of RAD51 or BRCA2 enhances SSA [80, 332, 333]. This regulation suggests that SSA operates as a backup pathway to HRR, especially during S-phase, although its actual role remains to be better-characterized [160].

3.2. Gene Targeting by Homology-Directed Repair (HDR)

As HRR naturally utilizes the sister chromatid as a template for DSB repair [89], the introduction of artificial nucleases, along with a targeting vector containing two homologous sequences (homology arms) to the break site, enables high efficiency of gene targeting mediated by HDR. Based on this, three different strategies could be used (**Figure 4**):

a) Gene correction: Adopted for the replacement of disease-causing mutations of few pairs of bases (point or small mutations) by the corrected pair of bases.

b) Knock-in: The therapeutic cDNA is inserted into the endogenous location of the disease-causing mutated gene. The endogenous elements of the locus will control and coordinate the expression of the exogenous cDNA. This strategy can be applicable to any kind of mutation in a determined gene.

c) *Safe harbor* integration: The whole therapeutic cassette, even including an exogenous promoter is introduced into a *safe harbor* locus. This strategy is very useful when the disease is produced by different mutations not only in one gene but in different genes.

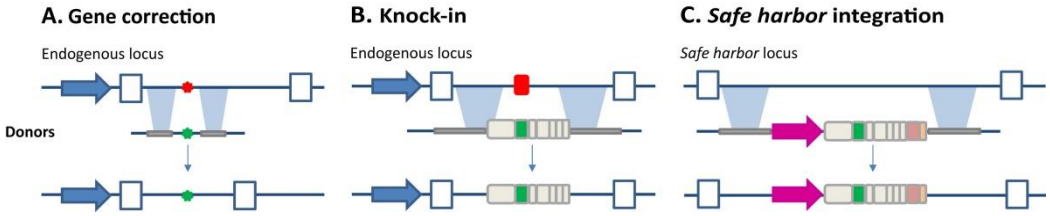


Figure 4. Gene-editing strategies by HDR. A) Gene correction. B) Knock-in. C) *Safe harbor* integration. The mutation is indicated in red, and when corrected, in green. The homology arms are indicated in grey, the endogenous promoter in blue, and the exogenous in pink. Modified from Garate et al. [172].

3.2.1. Genomic *Safe harbor* (GSH) loci

By definition, a genomic *safe harbor* locus is a genomic site able to accommodate the integration of new genetic material in such a way that the newly inserted genetic elements function in a predictable manner, facilitating enough transgene expression and without causing adverse effects on the host cell due to activation of adjacent genes or gene disruption [417, 468]. *Safe harbor* loci could be located both in intragenic and extragenic regions. Considering intragenic regions, housekeeping genes and non-essential genes have been proposed as candidates but their documented safety is still incomplete. Therefore, extragenic regions have been proposed as safer areas; however, in these regions the chromatin is usually very compacted, hampering gene targeting or silencing the expression of the transgene. Although

five GSH criteria were proposed [416] and proof of principle was provided for different *safe harbor* loci that accomplished the criteria in a disease model of β-thalassemia, these criteria might evolve, taking into account all the emerging information in genomics.

3.2.1.1. Specific loci used for targeted transgene integration

Although none of the loci used so far for targeted transgene addition satisfy all the requirements for *bona fide* GSH; three specific loci have been used: the human C-C chemokine receptor type 5 (*CCR5*) [318, 526], the human and mouse *ROSA26* [242], and the human adeno-associated virus site 1 *(AAVS1).*

The **human *AAVS1* locus**, located on chromosome 19, was discovered as the integration site of the AAV in the genome [277], that encodes for the protein phosphatase I regulatory subunit 12C (PPP1R12C). This locus is also known as the myosin-binding subunit 85 (*MBS85*) that plays a role in regulation of actin-myosin assembly [244]. A considerable part of the population presents antibodies against AAV serotypes, suggesting that they harbor AAV integrations without any associated pathology [212]. Several studies of gene targeting have shown stable and robust expression of the transgene in different cell types with no gross abnormalities or adverse effects [120, 217, 318, 398, 445, 498, 598]. However, in other studies, no faithful transgene expression was reported due to epigenetic regulatory mechanisms that act in this locus [400]. In addition, transgene integration into the *AAVS1* and consequent reduction of MBS85 expression and changes in myosin II-dependent cellular contractile force have been recently published [352, 353]. Thus, the safety and faithful expression in this locus remain to be validated.

3.2.1.2. The mouse ortholog locus of the human *AAVS1* locus

In this study, we have implemented a *safe harbor* integration strategy in a mouse locus orthologous to the human *AAVS1*. Gene targeting in this locus allows the characterization of the AAV site-specific integration site in a mouse model to determine if the disruption of this locus is safe [144]. The mouse *Mbs85* locus is located on chromosome 7, in a region that is syntenic to the human *AAVS1*. This gene spans 20 kilo-bases (kb), and the resulting 3.1 Kb mouse cDNA and protein are 77% and 86% identical to their human counterparts, respectively [522]. Moreover, the replication sequences (TRS/RBS motifs) needed for the replication and integration of the AAV are present in the mouse locus [212, 304, 305]. Mouse embryonic stem cells (mESCs), carrying a specific integration site of a recombinant AAV (rAAV), did not demonstrate any aberrant expression, even after injection into blastocysts *in vivo*. Moreover the EGFP expression encoded by the rAAV remained strong without altering the neighboring genes [211].

3.3. Artificial Programmable Nucleases

Four major platforms currently exist for inducing site-specific DSBs: meganucleases (MNs), zinc finger nucleases (ZFNs), transcription activator-like effector (TALE)-nucleases (TALENs), and more recently discovered the RNA-guided cluster regularly interspaced short palindromic repeats (CRISPR) system (**Figure 5**). In **Table 2** are indicated their main characteristics.

Platform	Natural origin	Advantages	Disadvantages	Target-binding principle & Recognition sequence & Configuration & Cleavage ends
MN [98, 121, 147, 491]	Present in all kingdoms, most of them encoded by introns and inteins (LAGLIDADG family being the most studied).	-Small size (1 Kb). -Poorly immunogenic in mammalian cells.	-Only cleave regions in which the restriction site of the enzyme is present. -They present reduced activity due to methylation of the targeted locus.	-Protein-DNA specific recognition based on different motifs. -Recognize 14-40 bp. -Monomers or homodimers. -Generation of 3' overhang.
ZFN [271, 272, 19]	ZF (zinc finger) motif is the most common DNA binding motif in eukaryotes.	Good efficiency and specificity.	-Laborious design, months. -Context dependency.	-Protein-DNA specific recognition based on 3-4 ZF that bind DNA (each ZF recognizes 3 bp) -Recognize 9-12 bp. -Dimer (>heterodimers). -5' overhang generated by *FokI* endonuclease.
TALEN [37]	TALE are bacterial proteins that are injected into infected plants via a type III secretion mechanism by *Xhantomonas* pathogens	-Easy design, takes 1-3 weeks. -Very good efficiency and specificity. -Minimal off-targets.	-T at the 5' recognition site. -High size (6 Kb). -Highly-repetitive sequence.	-Protein-DNA specific recognition based on 15.5-19.5 repeated TALE units binding DNA (each repeat unit recognizes 1 bp) -Recognized 10-30 bp. -Dimer (>heterodimers). -5' overhang generated by *FokI* endonuclease.
RNA-guided CRISPR system [24, 78, 100, 154, 228, 404, 151, 588]	Adaptive immune system that evolved in Bacteria and Archaea to battle against invading plasmids and viruses.	-Very easy design, 2-5 days. -High efficiency. -Small size (3.1-4.3 Kb). -Versatility to target different foci at the same time.	-PAM sequence restriction. -Most prone to generate off-targets.	-Watson-Crick rule with a short RNA molecule. -Recognize 20-22 bp. -Monomers. -Blunt ends (or overhangs depending on the nuclease).

Table 2. Summary of the main gene-editing platforms including their different characteristics.

3.3.1. Transcription Activator-Like Effector Nucleases (TALEN)

TALEN were generated by fusing the TALE motifs to the catalytic domain of the *FokI* endonuclease, working as a dimer to cleave the DNA target site [103, 350], separated by a DNA spacer of 10-30 bp [328] that allows the formation of the active dimer [130, 497]. The TALE DNA binding domain generally comprises a tandem array of 15.5-19.5 repeats, each one consisting of approximately 34 amino acids [36], the last being 0.5 repeat modules shorter [377]. The amino acid sequences of each repeat are nearly identical except for two adjacent amino acids (the repeat-variable di-residues or RVDs) in position 12 and 13 that give specificity. The fact that some RVDs are able to bind to different bases has allowed the generation of TALE arrays able to bind to different DNA sequences. Modifications on the natural TALE scaffold were generated on either site of the repeat units, and TALENs demonstrated comparable activity to the ZFN counterparts [350, 376]. Nowadays, many platforms exist for engineering TALE arrays, ranging from the simplest methods to the latest techniques [49, 233, 298, 369, 454, 472, 481, 566, 590].

In the present study, we have used two pairs of TALEN with different RVD specificities that target the *Mbs85* locus in primary cells, both mouse HSCs and MEFs.

3.3.2. Considerations for the application of gene-editing platforms

3.3.2.1. Specificity and off-target outline

High on-target efficiencies can be obtained with any of the systems, although off-target strand breaks may also be introduced. The dose and time exposure to nucleases are critical to producing DSBs favouring on-target specificity. However, the principal determinant is the length of the DNA binding sequence of the nucleases [265]. Theoretically, a 16 bp recognition site should already be statistically unique in the human genome, but less than 50% of the 16 bp combinations are unique [295, 312]. Thus, due to the length of their DNA recognition motif, TALENs are expected to bind fewer off-targets than other gene-editing tools [346].

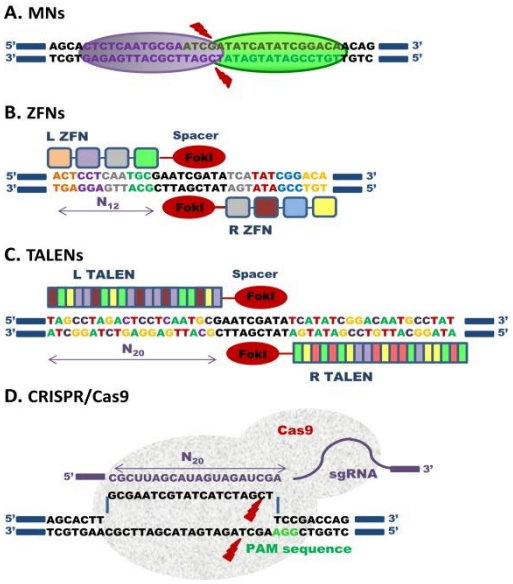


Figure 5. Gene-editing platforms. **A)** MNs; **B)** ZFNs; **C)** TALEN; **D)** CRISPR/Cas9. Read text for more details.

Furthermore, nucleases working as a dimer such as ZFNs or TALEN, or even CRISPR-dCas9 fused to the *FokI* domain, have been proved to produce fewer off-targets, because they were obligated to function as a dimer limiting their DNA binding possibilities [101, 447, 471]. In the CRISPR/Cas9 system, reduced off-target activity has also been observed by sgRNAs with truncated 3' ends (within the tracrRNA-derived sequence) or with two extra guanine nucleotides in the 5' end [166], by using paired nickases, or by using Cas9 variants [61, 227, 273, 330, 471]. Nowadays, new sophisticated techniques progressively arise in the analysis of off-targets [85, 213, 275] and in the future, whole genome sequencing in polyclonal populations with techniques with very high sensitivity rates would be desired to detect real off-targets [346, 539].

3.3.2.2. Nuclease and donor delivery

Nuclease and donor delivery remain one of the most significant challenges of genome editing methods. Artificial nucleases are delivered either as a plasmid DNA, *in vitro*-transcribed mRNA or as viral vectors; with respect to the donors, they are mostly introduced by means of plasmid DNA or via viral vectors.

The most widely used method for the introduction of the nucleases is transfection of plasmid DNA [431, 542]. The donor could also be introduced as a plasmid DNA, either as a circular [337, 486] or linear DNA [401] or single stranded oligonucleotides [92]. However, DNA transfection induces high cellular cytotoxicity and it is highly immunogenic due to the presence of bacterial sequences in the plasmid backbones. Additionally, random integration into the genome should be considered [359, 574]. Thus, transfection of mRNA [280, 357, 397, 469] or protein delivery of the nuclease, with a more limited duration of nuclease activity, have also been applied with a high percentage of gene editing [168, 269, 310, 446, 485, 599]. Furthermore, new self-assembled nanoparticles [514] could potentially be used in the near future [229]. Viral vehicles using integrase-deficient LV vectors (IDLVs), adenovirus vectors (AdVs) and adeno-associated viral vectors (AAVs)

constitute optimal tools for the delivery both of the donor or of the nucleases [94]. Apart from being delivery agents by themselves, they could also be used as a protein and mRNA transfer tools [54, 356]. AdVs do not integrate and have a large packaged capacity, and they have been used to deliver nucleases that target $CD4^+$ T cells [424] and HSPCs [477]. Concerning AAVs, despite their low packaging capacity, have demonstrated their efficacy in the delivery of nucleases [162, 169, 204, 390, 448] and donor [15, 113, 474, 560]. And finally, high efficiencies of transduction of HSPCs have been obtained with IDLVs (which do not integrate) delivering the donor template [184, 317].

3.4. Relevant targeted *ex-vivo* GT approaches in blood-related diseases

Of particular interest has been the *ex vivo* modification of $CD4^+$ T cells to knock out the *CCR5* coreceptor used for primary human immunodeficiency virus 1 (HIV-1) infection. Based on successful preclinical studies [221, 424], different clinical trials have been established (view summary in [325]), and one of them has already demonstrated the safety of the approach [526]. Moreover, among the studies performed in primary immunodeficiencies [434], the ones performed with ZFN to induce HDR into the *IL2RG* locus in human HSCs are remarkable[184, 317, 542]. This has also been performed in HSPCs from a CGD patient to induce HDR with ZFN into the *AAVS1* locus [113]. With respect to hemoglobinopathies, ZFNs have been used to repair the human β-globin gene in $CD34^+$ from sickle cell disease (SCD) patients [216]. Furthermore, CRISPR/Cas9, ZFNs and TALENs have been used in $CD34^+$ to generate by NHEJ disruption of the *BCL11A* enhancer region, obtaining fetal globin levels that could be therapeutic for SCD or β-thalassemia [58, 551]. Finally, different gene-editing platforms have been used to correct by HDR different immunodeficiencies, erythroid diseases, and hemophilia in patient-induced pluripotent stem cells (iPSCs) [131, 235, 321, 322, 345, 347, 418, 419, 486, 513, 578, 579].

3.5. Targeted GT approaches in Fanconi anemia

Five different studies of gene editing have been performed with FA cells. In 2009, a first study was published about the use of AAVs without nucleases to correct a B lymphoblastoid cell line of a FA-A patient with very reduced efficiencies of 0.016% HRR [412]. Our laboratory achieved the first results with nucleases, using AdV5/35-ZFNs that targeted the human *AAVS1* locus, in combination with an IDLV donor carrying the human *FANCA* gene in primary fibroblasts from FA-A patients. Corrected fibroblast cells (that reached up to 40% of correction due to their proliferation advantage) were reprogrammed to generate iPSCs that were subsequently differentiated into disease-free functional HPSCs [456]. A second study was performed in p53-downregulated iPSCs from FA-A fibroblasts that were gene-edited with helper-dependent AdVs, without using nucleases. This study used a knock-in strategy in the iPSCs that were finally differentiated into three different lineages [309]. Another study demonstrated the feasibility of performing HDR knock-in into the c.456+4A>T mutation to restore proper expression of FANCC gene in FA-C patient fibroblasts. The authors used the CRIPR/Cas9 platform either as a nuclease or as a nickase, with a donor template carrying the human FANCC gene. Higher gene correction frequencies were obtained with the Cas9 as a nickase, although with the two systems they obtained normalization of the FANCC expression gene [403]. Lastly, this group has recently showed 66-70% HDR (using Cas9 as nickase) in selected iPSCs derived from reprogrammed FANCI fibroblasts that were efficiently differentiated into hematopoietic progenitors [402].

VI. OBJECTIVES

The aim of this study was to correct the phenotype of hematopoietic stem and progenitor cells (HSPCs) from a mouse model of Fanconi anemia A (FA-A) by means of the targeted integration of the human *FANCA* gene into the mouse *Mbs85* locus, an ortholog of the human *AAVS1 safe harbor* locus.

To achieve this main goal, the following objectives were proposed:

1. To generate two donor constructs carrying either the therapeutic human *FANCA* gene or the *EGFP* reporter gene, respectively.

2. To demonstrate the cleavage efficacy of designed TALEN in the mouse *Mbs85* locus.

3. To develop a therapeutic gene-targeting approach to correct the phenotype of FA-A mouse embryonic fibroblasts.

4. To demonstrate the targeted integration of the donor constructs into the *Mbs85* locus of hematopoietic stem and progenitor cells from WT and FA-A mice.

5. To evaluate the therapeutic potential of the proposed gene-targeting approach in hematopoietic stem and progenitor cells from FA-A mice.

VII. MATERIALS AND METHODS

1. TALEN AND DONOR CONSTRUCTS

1.1. TALE nucleases and plasmids

Two pairs of TALEN were designed by Dr. Claudio Mussolino (Institute for Cell and Gene Therapy at University Medical Center in Freiburg, Germany) to target intron 1 of the *Mouse Myosin Binding Subunit 85* (*Mbs85*) locus, also known as the *Protein Phosphatase 1 Regulatory Subunit 12C* (*Ppp1r12c*), located in Chromosome 7: 4,481,520-4,501,680. The TALEN pairs generated to target this locus will be referred to during this work as mouse TALEN (mTALEN). These mTALEN specifically target the 5'-TGTCCTCTCTTCTTGCTAG NN NN AGTTACTGGTGGGAACAGA-3' sequence. Different spacer lengths were tested, and finally, the mTALEN were designed with a spacer of 13 bp between the two monomers.

The DNA binding domain of each mTALEN monomer comprised a tandem array of 17.5 single repeats, each one consisting of 34 highly conserved residues. The RVDs in the 12th and 13th amino acid positions of each repeat specify the DNA base being targeted according to the cipher NG = T, HD = C, NI = A, NK= G and NN = G or A. Two different mTALEN pairs were designed, characterized because they contain different amino acid modules that recognize guanine (G) nucleotides. The NN-mTALEN pair has NN modules that recognize a G or an adenine (A) nucleotide, while the NK-mTALEN pair has NK modules that recognize a G nucleotide.

The TALEN backbone of both mTALEN contains an N-terminal domain (which contains 153 residues) with a nuclear localization signal (NLS), the 0 repeat binding to the 5' T -nucleotide, the 17.5 repeats, and the C-terminal domain (which contains 17 residues) fused to the catalytic domain of the *FokI* endonuclease type II [377]. A scheme of the engineered mTALEN (specifically represented the NN-mTALEN pair) and the targeted site in the mouse *Mbs85* locus is described in **Figure 6 A**. The mTALEN backbone was cloned in the laboratory of Dr. Claudio Mussolino in a pVAX plasmid that contains a cytomegalovirus (CMV) promoter. Two different plasmids were generated, one for the left mTALEN monomer and another for the right mTALEN monomer. In both plasmids there is a T7 promoter that could drive the transcription of the TALEN, followed by hemagglutinin (HA) tag, which allows the monitoring of the TALEN expression, then the NLS, and the DNA binding repeats that recognize the left or right site of the *Mbs85* locus. The backbone also contains the *FokI* domain, and the *PmeI* restriction enzyme site to allow the linearization of the plasmid for *in vitro* mRNA transcription (**Figure 6 B**).

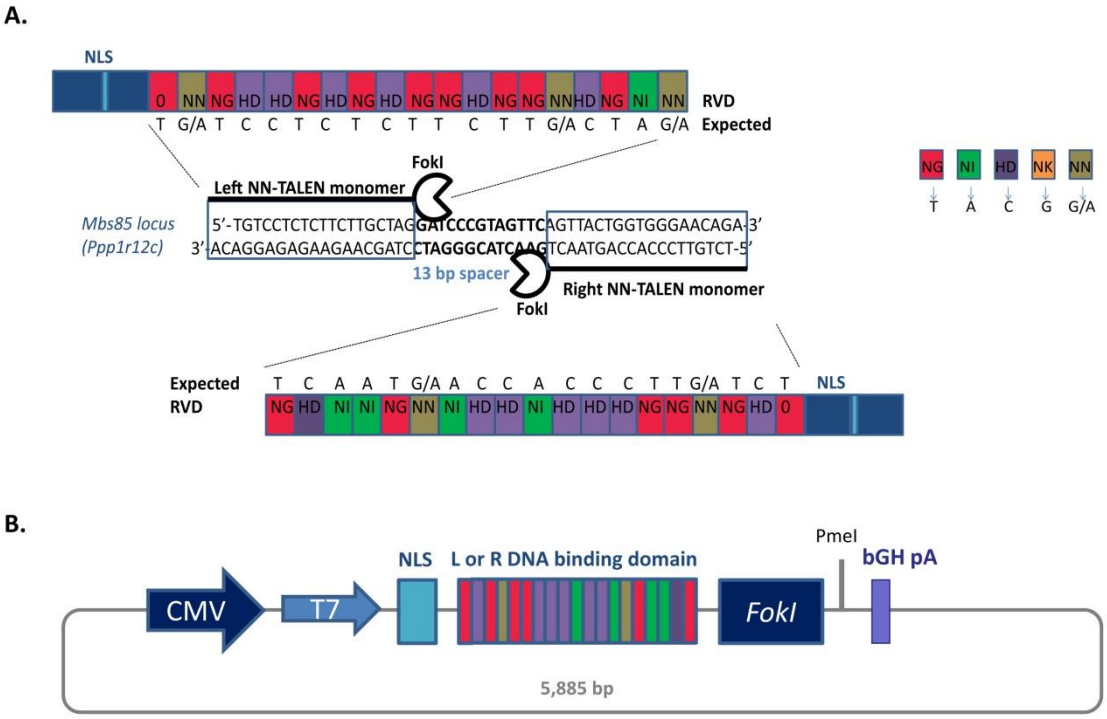


Figure 6. mTALEN design. **A)** Target sites of the left (L) and right (R) monomers of the NN-mTALEN in the *Mbs85* locus. The amino acids DNA binding modules of the engineered NN-mTALENs, as well as the expected target sequences are indicated. The N-terminal region that contains the nuclear localization signals (NLS) as well as the C-terminal region that contains the *FokI* cleavage domain are also indicated. The amino acids modules are represented with different colors based on the recognized nucleotides, thymine (T), adenine (A), cytosine (C) and guanine (G). **B)** Schematic representation of the mTALEN plasmid. CMV: CMV promoter; T7: T7 promoter from which the mRNA is synthesized *in vitro*; NLS: nuclear localization signal; L or R DNA binding domains; *FokI*: *FokI* cleavage domain; bGH pA: bovine growth hormone polyadenylation signal. *PmeI* indicates the restriction site present that allows linearization of the vector for *in vitro* mRNA synthesis. The total size of the plasmid is 5,885 bp.

1.2. Donor constructs

Donor templates were used to promote the repair of DNA lesions, such as the DSBs generated by the TALEN, by homologous directed repair (HDR). HDR uses long stretches of homologous sequences between the damaged genomic DNA and the donor construct to repair the DNA lesions. To favour the introduction of an external therapeutic cassette, the donor must contain two homology sequences at both sites of the integration, in the left and right sides. These two homologous sequences are also referred to as mouse homology arms (mHAs). In our particular case, the mHAs were homologous to the *Mbs85* locus, which was targeted by the mTALEN. We designed the left and the right HAs that contain a sequence of 806 bp and 860 bp, respectively.

1.2.1. PGK-h*FANCA* therapeutic donor plasmid

The therapeutic donor carries the physiological phosphoglycerate kinase 1 (PGK) promoter that drives the expression of the WT cDNA of the human *FANCA* gene. The safety of this promoter has already been demonstrated in GT studies [32, 189, 367]. This donor has a 3XFLAG signaling peptide in the N-terminus of the *FANCA* gene, a self-cleaving 2A peptide (E2A) sequence [520] and a puromycin resistance selection gene followed by a SV40 PolyA (Simian virus 40 PolyA) sequence. The whole cassette is flanked by each mHA. This plasmid has a size of 10,303 bp.

The therapeutic cassette was chemically synthesized by GenScript (Piscataway Township, USA) in a pUC57 backbone. The 3XFLAG signalling peptide added to the N-terminus of the *FANCA* gene was constructed by fusing 3 tandem FLAG epitopes of 22 amino acids in length. The detection sensitivity of fusion proteins containing 3XFLAG is significantly enhanced in mammalian cells [594]. Moreover, the risk of altering protein functions by blocking other epitopes, decreasing their solubility, or modifying their localization is minimal with this 3XFLAG sequence [580].

This therapeutic cassette was designed to easily remove any fragment in case it was necessary. Thus, different multicloning sites were located throughout the cassette. The donor with its multicloning sites is represented in **Figure 7 A**.

This donor has been used in gene-editing experiments, both in the studies performed in FA-A MEFs and those performed in FVB FA-A lineage negative (Lin^-) BM cells.

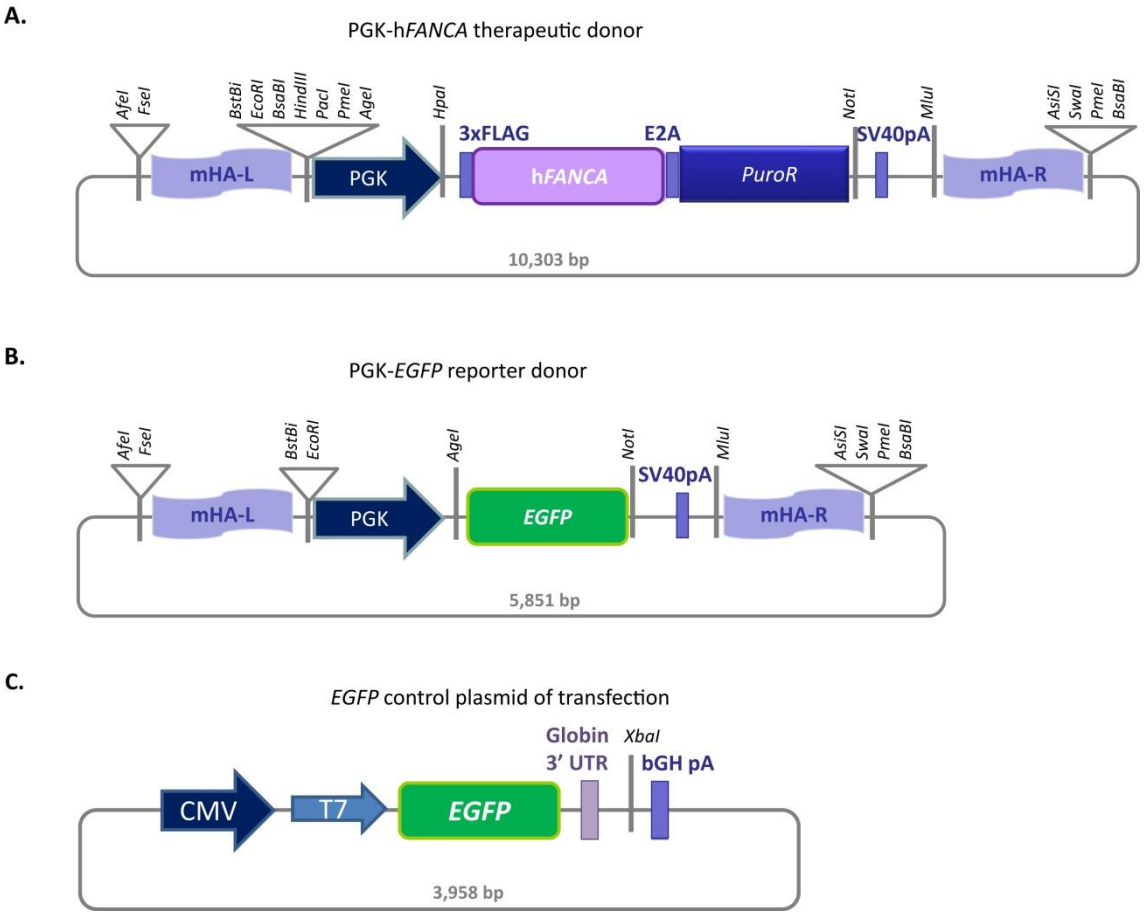


Figure 7. Schematic representation of the h*FANCA* and *EGFP* donor constructs, and the *EGFP* plasmid used as a control of transfection. **A)** PGK-h*FANCA* therapeutic donor. The plasmid includes the left (L) and right (R) mHAs, the PGK promoter, 3XFLAG signal bound to the human *FANCA* gene, the E2A sequence, the puromycin resistant gene (*PuroR*), and the simian virus polyadenylation signal (SV40 pA). The different multicloning sites are indicated with the names of the enzymes. The total size of the plasmid is 10,303 bp. **B)** The PGK-*EGFP* reporter donor contains the same backbone elements, although instead of the 3XFLAG-h*FANCA*-E2A-*PuroR* cassette it contains an *EGFP* reporter gene. The total size of the plasmid is 5,851 bp. **C)** The *EGFP* control plasmid used in transfections. CMV: CMV promoter; T7: T7 promoter from which the mRNA is synthesized *in vitro*; 3' untranslated region (UTR): β-globin 3'UTR regulatory sequence; bGH pA: bovine growth hormone polyadenylation signal. *XbaI* indicates the restriction site that allows linearization of the vector for *in vitro* mRNA synthesis. The total size of the plasmid is 3,958 bp.

1.2.2. PGK-*EGFP* reporter donor plasmid

A PGK-*EGFP* reporter donor was constructed in order to demonstrate targeted integration (TI) with a smaller donor that contains a fluorescent selection marker. This donor was used in gene-editing experiments performed in C57BL/6 WT Lin^- BM cells.

The expression of the *EGFP* was driven by the physiological PGK promoter, and the *EGFP* sequence was followed by a SV40 PolyA. The whole cassette was flanked by each mHA that were the same as the ones of the therapeutic PGK-h*FANCA* donor. This plasmid has a size of 5,851 bp. The plasmid is represented in **Figure 7 B.**

This plasmid was cloned in our laboratory at CIEMAT. Briefly, both the therapeutic *FANCA* plasmid (that contains the main backbone with the corresponding mHA and the SV40 PolyA sequence) and the pCCL.PGK.*EGFP*.wPRE* plasmid (that contains the PGK.*EGFP* fragment) were digested with *EcoRI*/ *NotI* restriction enzymes (New England Biolabs, Ipswich, Massachusetts, USA). Electrophoresis was then performed to separate the expected bands of 4,857 bp and 1,264 bp corresponding to the vector backbone and to the insert. Later on, these bands were purified with the NucleoSpin® Gel and PCR Clean-up kit (Macherey Nagel, Düren, Germany). After that, the purified fragments were ligated with the T4 ligase (New England Biolabs) and the resultant products transformed in TOP10 bacteria (Invitrogen, Life Technologies, Thermo Fisher Scientific) and plated in Ampiciline resistant plates. Plasmid DNA was purified from bacteria mini-cultures (generated from picked colonies) using the Fast Plasmid Mini Kit (5 Prime, Hilden, Germany) and next, it was analysed with different pairs of enzymes *EcoRI*/*PmeI*, *EcoRI*/*NcoI*, and *EcoRI*/*XbaI* (New England Biolabs) to confirm the correct ligation. Finally an endo-free MAXIprep (NucleoBond® Xtra Midi EF, Macherey Nagel) was generated to use this plasmid for transfection experiments.

1.3. Control plasmids

1.3.1. PGK-h*FANCA* control plasmid of targeted integration

In order to have a positive control to tune up the polymerase chain reaction (PCR) conditions for targeted integration, the following plasmid was designed and extended by chemical synthesis by GenScript (Piscataway Township, USA). Its sequence was exactly the same as that of the therapeutic PGK-h*FANCA* donor except for the mHAs that were 300 bp longer towards the left and right ends, thus including longer parts of the left and right *Mbs85* locus.

1.3.2. *EGFP* control plasmid for transfections

This plasmid was used in all the transfections, both lipofections and nucleofections as a control of transfection efficiency. This *EGFP* control plasmid was constructed in the pVAX backbone that is the same one as that of the mTALEN, and thus, it has the same elements intrinsic to the backbone of the plasmid, and reflects what happens with the transfection of the mTALEN. The only difference is that the *EGFP* control plasmid includes a β-globin 3'UTR regulatory sequence for producing mRNA in a more stable manner [585] (**Figure 7 C**). This plasmid has a size of 3,958 bp. The mRNA generated from this plasmid will be referred to as mRNA *EGFP* (A).

Two other *EGFP* control plasmids were used in mRNA transfections. The first one contains the same backbone and elements as the above-mentioned plasmid, and it differs from the previous plasmid because it does not contain the 3'UTR regulatory sequence. The mRNA produced from this plasmid will be referred to as mRNA *EGFP* (B). This plasmid has a size of 3,700 bp. The second one is a commercial StemMACS Nuclear *EGFP* mRNA (MACS, Miltenyi Biotec, Bergisch Gladbach, Germany) which will be referred to as mRNA *EGFP* (C). This mRNA is an extremely pure nuclear *EGFP* mRNA recommended as a positive control for mRNA transfections of nuclear factors.

The pMAX *EGFP*® vector which is provided by Lonza kits and has a size of 3,486 bp was also used as plasmid DNA in some nucleofection experiments.

1.4. mRNA synthesis

EGFP control plasmids (A and B) (40 µg of DNA) were digested with *XbaI* restriction enzyme in order to linearize the plasmid. This digestion was purified using phenol-chloroform and the amount of DNA was measured with a spectrophotometer (Nanodrop® ND.1000, Thermo Fisher Scientific). Then, the *in vitro* synthesis of the mRNA was performed with the MegaScript T7 Kit (Ambion/Life Technologies/Thermo Fisher Scientific) and 3'-O-Me-M7G (5')ppp(5')G RNA Cap Structure Analog (ARCA) (New England Biolabs). First, 6 µg of the digested DNA were mixed with 32 µl of dNTPs, 12 µl of ARCA, 8 µL of T7 enzyme, 8 µL of T7 Buffer 10X and up to 80 µL with nuclease-free water and incubated for two hours (h) at 37°C. Next, 4 µL of Turbo DNase were added to the reaction and incubated again for 30 minutes at 37°C. Next, the polyadenylation reaction was performed in the synthesized RNA using the Poly(A) Tailing Kit (Ambion). For the RNA reaction synthesis, 144 µL of nuclease-free water, 80 µL EPAP buffer 5X, 40 µL $MnCl_2$ 25 mM, 40 µL ATP 10 mM and 16 µL EPAP enzyme were added, mixed and incubated for one hour at 37°C.

The resulting mRNA was purified using RNeasy Plus Mini (Qiagen, Venlo, Limburg, Netherlands) following the manufacturer's protocol. Finally, the amount of mRNA was measured with Nanodrop®, aliquoted and stored at -80°C. An aliquot of the synthesized mRNA was run in an agarose gel in denaturing conditions (with formaldehyde) in order to confirm the integrity of the mRNA before and after the polyadenylation reaction.

The different doses and conditions used to transfect the mRNA are explained in the different protocols and experiments. The different quantities used during the experiments have been calculated in picomoles of mRNA, as the number of molecules transfected could vary if we use DNA or mRNA (**Table 3** summarizes the doses frequently used of the different constructs as DNA or mRNA).

mRNA constructs	Size of the nucl. acid [pb]	µg	pmol
***EGFP* Batch A**	939 + 150 (pA)=1,089	0.4 2 4 5 10 15 25 50	1.08 5.39 10.79 13.48 26.97 40.45 67.42 134.84
***EGFP* Batch B**	813 + 150 (pA)=963	3 4 6 12	9.15 12.2 18.3 36.6
DNA plasmids	**Size of the nucl. acid [pb]**	**µg**	**pmol**
mTALEN	5,885	0.4 0.75 1 1.5 2 2.5 3	0.11 0.21 0.26 0.39 0.52 0.65 0.79
Donor PGK-h*FANCA*	10,303	0.75 2 2.5 4 5	0.11 0.30 0.37 0.60 0.75
Donor PGK-*EGFP*	5,851	2 4	0.53 1.05
***EGFP* Batch A**	3,958	0.4 2 2.5	0.16 0.78 0.97

Table 3. Conversion: weight-picomoles (for nucleic acids). The different doses in micrograms (µg) of each mRNA or DNA constructs used throughout the experiments are also calculated in picomoles (pmol).

2. ANIMAL EXPERIMENTATION

2.1. General characteristics

All experimental procedures were carried out according to European and Spanish regulations (European convention ETS 123, regarding the use and protection of vertebrate mammals used in experimentation and other scientific purposes, Directive 2010/63/UE and Spanish Law 6/2013 and Real Decreto Law (R.D.) 53/2013 regarding the protection and use of animals in scientific research). Procedures involving Genetically Modified Organisms were carried out according to European and Spanish regulations (Directive 2009/41/CE and, Spanish Law 9/2003 and RD 178/2004). Procedures were approved by our Animal Experimentation Ethical Committee according to all external and internal bio-safety and bio-ethics guidelines, and authorized by the Spanish Government (Code PROEX # 070-15# about Cell and Gene Therapy in rare diseases with chromosomal instability).

Mice were housed and bred at the CIEMAT Laboratory Animals Facility (registration number ES280790000183). All mice were routinely screened for pathogens in accordance with the Spanish Society for the Laboratory Animal Science (SECAL) and the Federation of European Laboratory Animal Science Associations (FELASA, Tomworth, United Kingdom) recommendations and no pathogens were found.

Mice were provided with food (TEKLAD Global Diet 2918, irradiated with 25 KGy gamma rays) and water (50 µm filtered and UV irradiated) *ad libitum*, under controlled environmental conditions. Mice were housed during the experimental protocols in micro insulators individually ventilated cages type IIL with 25 air cage changes per hour. A maximum of 6 mice were housed in each cage. Room lighting was controlled with light/dark cycles of 13/11 hours, and temperature and humidity were regulated at 20 ± 2ºC and 55 ± 10%, respectively. HEPA air filters were present in all rooms.

The department of Basic Research at CIEMAT has a category 2 radioactive installation which houses an X-ray device. It is mainly used for the irradiation of laboratory animals, biological samples and other materials, according to Spanish laws such as R.D. 35/2008 regarding the regulation on nuclear and radioactive facilities and R.D. 783/2001 regarding the regulation on health protection against ionizing radiation.

2.2. Mouse strains

C57BL/6J, also known as B6 (B6CD45.2), were used as WT controls and donors in most of the experiments, and P3B (B6.SJL-Ptprc^{a} Pepc^{b}/BoyJ) mice, also known as B6CD45.1 (congenic strain that carries the differential pan-leukocyte marker designated CD45.1 (Ptprc^{a} allele)), were used as recipient mice in transplant experiments. The colony founders of both strains were obtained from Jackson Laboratory Animals (Bar Harbor, Maine, U.S.A.). Experimentation animals of both strains of 8-14 weeks-old were used in the majority of experiments.

FVB/NJ mice (from Envigo, Barcelona, Spain) also known as FVB were used as WT controls in the experiments involving the use of *Fanca*$^{-/-}$ mice; whereas FVB/*Fanca* $^{\Delta E4\text{-}E7.Ins.LacZNeo}$ (referred to as *Fanca*$^{-/-}$ or FA-A) were used as a mouse model of Fanconi Anemia. These mice have a neomycin resistance cassette that replaces exons 4-7 of the gene *FANCA*. The founders

of the colony, which now are in the Laboratory Animal Facility at CIEMAT, were obtained from Dr. Fre Arwert, Department of Clinical Genetics and Human Genetics at the Free University, Amsterdam (Netherlands). Animals of both strains from 8-14 weeks-old were used.

The characteristics of the mouse strains used during the experiments are summarized in **Table 4**.

Strain	Type	Background	H2	CD45	*Fanca* ΔE4-E7	Donor/ Recipient	Experiments
B6 or C57BL/6J	Inbred	C57BL/6J	b/b	CD45.2/ CD45.2	-	Donor	*In vitro* and *in vivo* gene-targeting experiments
P3B or B6.SJL-*Ptprc*^{a} *Pepc*^{b}*/BoyJ*	Inbred	C57BL/6J	b/b	CD45.1/ CD45.1	-	Recipient	*In vivo* gene-targeting experiments
FVB NJ	Inbred	FVB/NJ	q/q	CD45.1/ CD45.1	-	Donor	*In vitro* gene-targeting experiments
FVB *Fanca* $^{\Delta E4\text{-}E7.Ins.LacZNeo}$	Inbred. Targeted mutation	FVB/NJ	q/q	CD45.1/ CD45.1	+	Donor/ Recipient	*In vitro* and *in vivo* gene-targeting experiments

Table 4. Characteristics of the mouse strains used throughout the thesis.

3. CELL LINES AND PRIMARY CELLS

3.1. Cell lines

The **293T** cell line (originally referred to as 293tsA1609*neo*) also known as HEK-293T, is a cell line derived from human embryonic kidney cells that has been modified and contains the simian virus SV40 T-antigen that is constitutively expressed, providing higher transfection efficiencies. The inserted sequence allows the episomal replication of plasmids that contain the SV40 origin of replication origin (ORI) [422].

The **NIH/3T3** cell line is an embryonic origin cell line from mouse sarcoma derived from the NIH/Swiss strain. 3T3 refers to the cell transfer and inoculation protocol for the immortalization of the line, and means "3-day transfer, inoculum $3x10^5$ cells". This cell line is highly sensitive to sarcoma virus focus formation and leukemia virus propagation, and has been proven to be very useful in DNA transfection studies.

The **Ba/F3** cell line is an interleukin-3 (IL-3) dependent murine pro B cell line. Cells were derived from a C3H mouse. Its growth and proliferation depends on the presence of IL-3, and cells undergo cell death by apoptosis upon factor withdrawal [413, 414]. This cell line was chosen because it is a blood-related cell line easy to nucleofect and it grows in suspension like lineage negative hematopoietic progenitors cells, which were our final goal.

The different cell lines used, their characteristics, and the culture conditions used in the present work are summarized in **Table 5**.

3.2. Primary cells

3.2.1. Mouse embryonic fibroblasts (WT and FA-A)

Mouse embryonic fibroblasts, both from FVB/NJ WT and FVB FA-A mice, were obtained from the chorion of 13.5 E (mouse embryonic day) pregnant females. Once the embryos were obtained from the uterus, they were separated from extra-embryonic tissues, and the head and viscera were removed. Then, the embryos were dissociated by mechanical and enzymatic protocols (with 0.05% Trypsin, Gibco/Life Technologies/Thermo Fisher, Waltham, USA). The large fragments remaining after dissociation were discarded by centrifugation and fibroblasts were seeded at a concentration equivalent to two embryos per 150 mm untreated plate (Corning, NY, USA) with the medium and under the conditions indicated in **Table 5**.

3.2.2. Immortalization of mouse embryonic fibroblasts (FA-A and WT)

MEFs from WT or FA-A mice were immortalized with a retrovirus that encodes for papillomavirus *16 E6-E7* genes. Retroviruses were generated by a transient transfection in HEK-293T cells with the transfer plasmid (pLXSN 16 *E6E7*), and a plasmid containing the gag-pol-env genes (pCL-ECO-gag-pol). Transfection was conducted in HEK-293T cells at 70% confluence in 100 mm plates (Corning, NY, USA) by the $CaCl_2$ DNA precipitation method.

Culture medium (IMDM (Gibco/Life Technologies/Thermo Fisher Scientific) with 10% Hyclone (GE Healthcare), 1% GlutaMAX™ (Gibco/Life Technologies/Thermo Fisher Scientific) and 0.5% Penicillin/Streptomycin (P/S) (Gibco/Life Technologies/Thermo Fisher Scientific) was replaced with fresh medium two hours before transfection. The amount of plasmid used for a 100 mm plate of HEK-293T cells was: 5 µg of the corresponding transfer plasmid (pLXSN 16 *E6E7*), and 5 µg of the gag-pol-env plasmid (pCL-ECO-gag-pol) [388]. This mixture was prepared in a final volume of 500 µl of 0.1X Tris-EDTA buffer/dH_2O (2:1) per plate and then 61 µl of 2.5M $CaCl_2$ was added. Next, the mixture was maintained for 15 minutes in agitation at room temperature (RT) to obtain a correct homogenization of the mixture. After that, 500 µl of 2X HBS buffer (100 mM HEPES, 281 mM NaCl, 1.5 mM Na_2HPO_4, pH 7.15) was added in drops while vortexing at full speed, allowing the formation of Ca^{2+}/DNA^- precipitates. After 13 hours, culture medium was replaced by fresh medium. Retroviral supernatants were collected 36 hours post-transfection and filtered through 0.2 µm pore-size filters (Millipore/Merck KGaA, Darmstadt, Germany). Viral supernatants were aliquoted and stored at -80ºC. The whole supernatant from two 100 mm plates was used to transduce (in two independent cycles of transduction) $6x10^6$ MEFs.

After 20 passages in culture, we considered that the immortalization of these cells was established. The protocol followed to immortalize the cells is described in **Figure 8**.

Immortalized MEFs were cultured in the same conditions as their non-immortalized counterparts (**Table 5**). Unless otherwise stated, immortalized fibroblasts have been used in most of the experiments performed with MEFs.

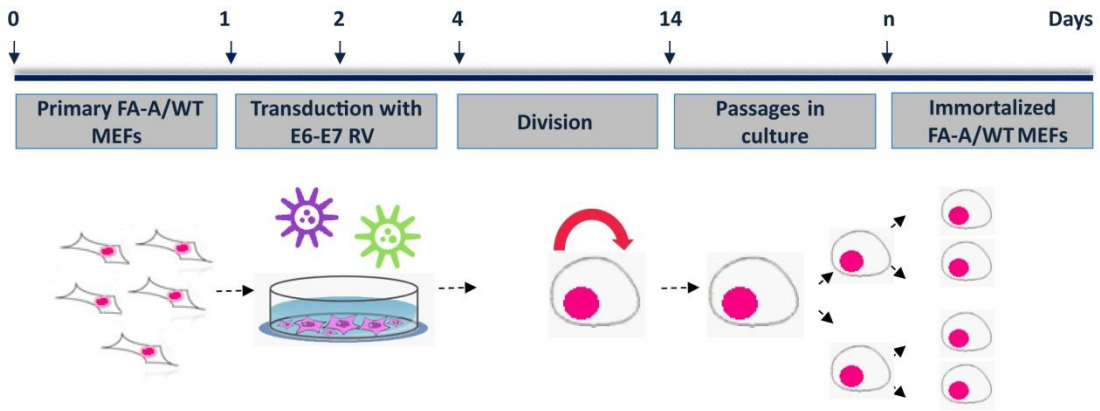


Figure 8. Scheme of the immortalization process of FA-A or WT MEFs, by transduction with a retroviral vector containing the HPV-16 *E6* and *E7* genes.

3.2.3. Mouse bone marrow cells

Bone marrow cells were isolated by flushing the femurs and tibias of these mice in IMDM (Gibco/Life Technologies/Thermo Fisher Scientific) with needles of 0.5x16 mm (25 G 5/3, BD Microlance 3). The cellular suspension was incubated with lysis solution ($ClNH_4$ with CO_3HK 1M with EDTA 0.5M) (Merck KGaA, Darmstadt, Germany) for 5 minutes at RT in darkness. Then, cells were washed with phosphate-buffered saline (PBS) 1X (Sigma® Life Sciences, San Luis, Missouri, USA) with 5% Hyclone (GE Healthcare) and 5% P/E (Gibco/Life Technologies/Thermo Fisher Scientific), and counted in a Neubauer chamber by diluting the sample in Türk solution (Merck). Lin^- cells were obtained by discarding mature hematopoietic cells, such as T cells, B cells, monocytes/ macrophages, granulocytes and erythrocytes and their committed precursors from the BM. The purification was done by immunoselection (see details in Materials & Methods section 4.1.1).

Cells	Organism	Cell type	Tissue & Disease	Growth	Origin	Culture medium	Culture conditions	Aim
HEK-293T	Human	Epithelial	Embryo Kidney/ No disease	Adherence	ATCC - CRL-3216™	Dulbecco's Modified Eagle's Medium (DMEM) (1X) with GlutaMAX™ (Gibco/LifeTechnologies/ ThermoFisher Scientific, Waltham, USA) or Iscove's Modified Dulbecco's Medium (IMDM, Gibco/Life Technologies/ThermoFisher Scientific) with GlutaMAX™, with 10% Hyclone (GE Healthcare) with 1% Penicillin/Streptomycin (P/S) (Gibco/Life Technologies/ThermoFisher).	Normoxia: 37ºC, 21% O_2, 5% CO_2 and 95% relative humidity (Thermo Scientific incubator, Whatman, USA).	-Production of RV -mTALEN expression
NIH/ 3T3	Mouse	Epithelial	Embryo/ Sarcoma	Adherence	ATCC® CRL-1658™	DMEM (1X) with GlutaMAX™ with 10% Hyclone with 0.5 % P/S. The culture plates were treated with 0.1% gelatin.	Hypoxia: 37ºC, 5% O_2, 5% CO_2 and 95% relative humidity.	mTALEN activity
BA/ F3	Mouse	IL-3 dependent murine pro B cell	Lymphatic background / No disease	Suspension	Provided by Dr. Toni Cathomen (DSMZ-ACC 300).	RPMI Medium 1640 (1X) with GlutaMAX™ with 10% Hyclone with 1% P/S with 10 ng/ml mouse interleukin-3 (mIL-3) (Genzyme, Cambridge, Massachusetts, USA).	Hypoxia: 37ºC, 5% O_2, 5% CO_2 and 95% relative humidity.	mTALEN activity
MEFs	Mouse	Epithelial	Embryo 13.5 E/ From FA-A or WT mice	Adherence	Primary or immortalized	DMEM with GlutaMAX™ with 10-20% Hyclone with 0.5% P/S. The culture plates were treated with 0.1% gelatin.	Hypoxia: 37ºC, 5% O_2, 5% CO_2 and 95% relative humidity.	*In vitro* gene-targeting experiments
Lin^- cells	Mouse	Hema-topoietic stem and progenitor cell	Bone Marrow/ From FA-A or WT mice	Suspension	Primary	In liquid culture they were maintained in IMDM (Gibco/Life Technologies/ThermoFisher Scientific) with 20% Hyclone, or StemSpan™ (StemCell™ Technologies, Vancouver, Canada) with 1% GlutaMAX™ with growth factors and cytokines and 1% P/S. The following factors were used unless otherwise stated: 100 ng/ml mouse stem cell factor (mSCF), 100 ng/ml human interleukin 11 (hIL-11), 100 ng/ml human FMS-like tyrosine kinase 3 ligand (hFlt3), 100 ng/ml human thrombopoietin (hTPO)(EuroBioSciences, Germany).	Hypoxia: 37ºC, 5% O_2, 5% CO_2 and 95% relative humidity.	*In vitro* and *in vivo* gene-targeting experiments

Table 5. Characteristics of the different cell types used throughout the thesis and their culture conditions.

4. CELL SORTING AND FLOW CYTOMETRY ANALYSES

4.1. Cell sorting

4.1.1. Purification of hematopoietic progenitors by cell sorting

Bone marrow cells were stained with a cocktail of antibodies against antigens for the lineage positive population (B220 (CD45R), Mac-1 (CD11b), Anti-Gr1 (Ly6G/C), CD3-ε, and Tert-119 antibodies conjugated with phycoerythrin (PE) (BD Pharmingen, San Jose, California, USA)) for 30 min at 4ºC. Of each antibody 0.3 µl were used, except in the case of Tert-119, in which 0.1 µl were used to stain one million cells in a final volume of 0.5-1 ml (depending on the total number of cells). More details of the corresponding antibodies are explained in **Table 6.** Once stained, cells were washed and resuspended in PBS 1X with 0.5 % Hyclone and with 5% P/E. Then, the lineage negative population (Lin^- BM cells) was purified by cell sorting (Cell sorter BD Influx, BD Bioscience).

MONOCLONAL ANTIBODIES					
Antigen	Conjugated	Clone	Host and Isotype	Stock Concentration (mg/ml)	Manufacturer
B220 (CD45R)	PE	RA3-6B2	Rat IgG2a, κ	0.5	BD Pharmingen
B220 (CD45R)	PE-Cy5	RA3-6B2	Rat IgG2a, κ	0.2	BD Pharmingen
CD3-ε	PE	145-2C11	Hamster IgG	0.2	BD Pharmingen
CD3-ε	PE-Cy5	145-2C11	Armenian Hamster IgG, κ	0.2	BD Pharmingen
CD3-ε	FITC	145-2C11	Hamster IgG	0.2	BD Pharmingen
Gr1 (Ly6G/C)	PE	RB6-8C5	Rat IgG2a, κ	0.2	BD Pharmingen
Mac-1 (CD11b)	PE	M1/70	Rat IgG2b	0.2	BD Pharmingen
Tert-119	PE	TER119	Rat IgG2a, κ	0.2	BD Pharmingen
Sca-1 (Ly6A/E)	APC-Cy7	D7	Rat IgG2a, κ	0.2	BioLegend
c-Kit (CD117)	A647	2B8	Rat IgG2a, κ	0.5	Southern
CD45.1 (Ly5.1)	PE	A20	Mouse IgG2b, κ	0.2	BD Pharmingen
CD45.2 (Ly5.2)	FITC	104	Mouse IgG2b, κ	0.5	BD Pharmingen
CD45.2 (Ly5.2)	Biotin	104	Mouse IgG2b, κ	0.5	BioLegend
SECONDAY ANTIBODIES					
Product	Stock Concentration (mg/ml)	Manufacturer			
SAV-PE-Cy7	0.2 µg/ml	Caltag			

Table 6. Flow cytometry antibodies used in the different experiments. Table indicates the antibody name, the conjugated fluorochrome, the clone where they were produced, the host and isotype, the stock concentration, as well as the manufacturer company.

4.1.2. Purification of CD45.2$^+$ population by cell sorting

Total BM cells were extracted from mice from experiments Lin$^-$ 29, 30, 32 and 34 (P3B recipients) and a cell sorting was performed to purify CD45.2$^+$ population. With this purpose, total BM cells were stained with 0.5 µl per each 10^6 cells of CD45.2$^+$ antibody conjugated with fluorescein isotiocyanate (FITC), and 0.5 µl per each 10^6 cells of CD45.1$^+$ conjugated with phycoerythrin (PE) (BD Pharmingen) (for more details about the antibodies see **Table 6**).

4.2. Flow cytometry analyses

In general, for flow cytometry analyses, also known as fluorescence-activated cell sorting (FACS) analyses, $5x10^4$ cells were collected and resuspended in 100 µl of flow cytometry buffer, PBA (PBS 1X with 0.5% bovine serum albumin (BSA, Fraction V; Sigma® Life Sciences) with 0.05% NaN_3) and stained for 30 minutes at 4°C with the corresponding primary antibodies (**Table 6**). Cells were then washed and resuspended in flow cytometry buffer containing 1 µg/ml of 4',6-diamidino-2-phenylindole (DAPI) (used as a viability marker) in PBA. Analyses were performed in the LSR Fortessa cell analyser (BD/ Becton, Dickinson and Company, New Jersey, USA). Off-line analyses were conducted with the FlowJo Software v7.6.5 (© FlowJo, LLC, Ashland, USA).

4.2.1. Analyses of viability and EGFP expression

These analyses were performed in different cell types. The viability was measured within the whole population (percentage of DAPI$^-$ cells), and the expression of EGFP was measured in the live-cell population (percentage of EGFP$^+$ DAPI$^-$ cells). Three different manners of gating the EGFP$^+$ population were carried out: SSC-A versus EGFP (also referred to as EGFP-1), EGFP versus FSC-A (also referred to as EGFP-2) and EGFP versus autofluorescence (also referred to as EGFP-3). An example of the gating strategy in FA-A MEFs is shown in **Figure 9 A**. This last gating was used to exclude the autofluorescence of the cells because cells' autofluorescence is detected both in EGFP (530/30) and 576/26 channels (**Figure 9 B**). The use of these three gating strategies to define the EGFP$^+$ population led us to corroborate the true EGFP$^+$ population.

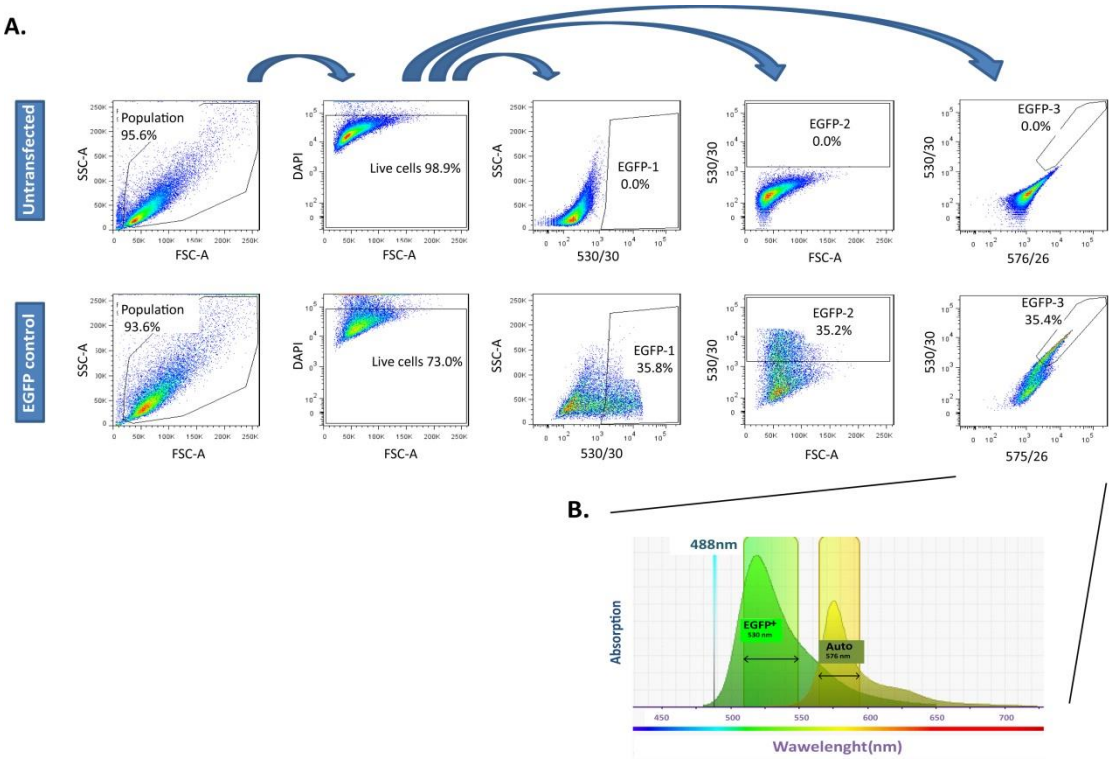


Figure 9. Gating strategy for the analysis of viability and EGFP expression by flow cytometry. A) Representative dot plot gating the whole population, the live-cell population after gating of singlets, and the EGFP$^+$ population with different gating strategies in FA-A MEFs, with untransfected cells and EGFP$^+$ transfected cells with a control plasmid. **B)** Spectral profile of the light emission of *EGFP* molecule in the 530/30 channel (green) and that of the autofluorescence from the cells that emits in the 576/26 channel (light yellow). The blue laser (488 nm) is used to excite the *EGFP* fluorophore.

4.2.2. Characterization of viability and transfection efficiency of Lin$^-$ cells and LSKs progenitors

Mouse HSPCs (mHSPCs) are Lin$^-$ cells that also express high levels of the stem cell antigen 1 (Sca-1) and receptor tyrosine kinase c-Kit (c-Kit) markers. These precursor cells are called LSK cells (lineage negative, Sca-1 positive, c-Kit positive) [240, 370, 504]. For their identification, Lin$^-$ cells were stained with 0.4 µl of a cocktail of antibodies against antigens for the lineage positive population (CD45R (B220), Mac-1 (CD11b), Anti-Gr1 (Ly- 6G/C), CD3-ε, and Tert-119 (0.2 µl) conjugated with phycoerythrin (PE) (BD Pharmingen), and also with 1.4 µl of the Sca-1 antibody conjugated with allophycocyanin-cyanine 7 (APC-Cy7) (BioLegend, San Diego, USA) and with 1.4 µl of the c-Kit antibody conjugated with Alexa Fluor® 647 dye (A647) (Southern, Atlanta, USA) (for more details see **Table 6**). Untransfected total BM cells and mock Lin$^-$ BM cells (transfected without DNA) were used as controls for the individual staining of these antibodies. Untransfected total BM cells were also stained with the CD3-ε primary antibody conjugated with FITC (BD Pharmingen) as a positive control of EGFP$^+$ cells.

After the nucleofection, viability was measured within the whole population (percentage of DAPI$^-$ cells), and the transfection efficiency was measured in the live-cell population (percentage of EGFP$^+$ DAPI$^-$ cells). Furthermore, the Lin$^-$ content and the LSK population were measured. Additionally, the EGFP$^+$ cells in the LSK population and in the whole population were gated and calculated. The gating strategy is represented in **Figure 10.**

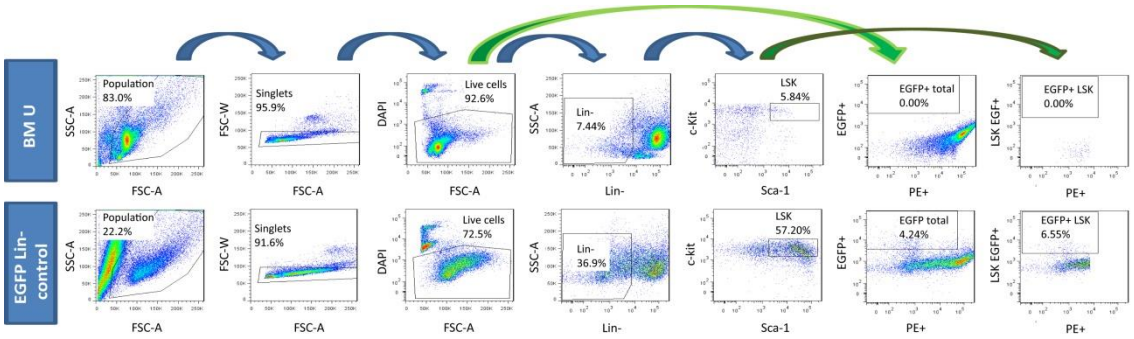


Figure 10. Gating strategy to characterize the nucleofected BM hematopoietic stem and progenitor cells in terms of viability and transfection efficiency. Representative dot plot gating the whole population, the singlets, the live-cell population, the Lin^- cells, the LSK population and the $EGFP^+$ population both in the whole population and in the LSK population. The represented gating strategy was performed in untransfected total BM cells as a negative control of EGFP expression, and in nucleofected $EGFP^+$ Lin^- cells as a positive control (nucleofected with an EGFP control plasmid).

5. HEMATOPOIETIC ASSAYS

5.1. Clonogenic assays of hematopoietic progenitors in methylcellulose

Clonogenic assays, also referred to as colony forming cell (CFC) assays, are *in vitro* experiments used to evaluate hematopoietic progenitor content. The analysis is based on the ability of hematopoietic progenitors to proliferate and differentiate into colonies in a semi-solid medium (methylcellulose) in response to cytokine stimulation.

Unless otherwise stated, $5x10^3$ Lin^- cells were seeded with methylcellulose-based medium containing mouse stem cell factor (mSCF), mouse interleukin-3 (mIL-3), and human interleukin-6 (hIL-6)(MethoCult™ M3534, StemCell Technologies, Vancouver, Canada). Each ml was seeded in triplicate (also referred to as experimental replicates) in treated 35 mm plates (Corning). The number of colony forming cells (CFCs) containing at least 50 cells was scored seven days later in an inverted microscope (for more details about the microscopes, see Materials & Methods section 12). All the colonies were maintained at 37ºC, in hypoxia with 5% O_2- 5% CO_2 and 95% relative humidity.

In order to analyse the number of gene-edited cells, $1x10^4$ or $2x10^4$ Lin^- cells were seeded per each 35 mm plate with MMC (at 10 and 30 nM) (Sigma® Life Sciences) or with puromycin (1 µg/ml) (Sigma® Life Sciences). The cells were seeded normally at 48 hours post-nucleofection with methylcellulose-based medium (MethoCult™ M3534, StemCell Technologies) and the corresponding dose of selection drugs. Triplicates of each condition were seeded in treated 35 mm plates (Corning) and culture in the same conditions as untreated cells (without the selection drugs). The number of colonies was scored seven days later in an inverted microscope or/and an inverted fluorescence microscope to analyse $EGFP^+$ colonies (see Materials & Methods section 12). Individual colonies were picked to analyse the targeted integration of the donor cassette into the *Mbs85* locus.

5.2. Hematopoietic transplantation assays

Recipient mice (P3B or FVB FA-A) were irradiated with myeloablative conditioning of 9 Gy, two doses of 4.5 Gy separated 24 hours, with an X-ray equipment MG324 (Philips, Hamburg, Germany), at 300 kV and 10mA, with a dose rate of 1.03 Gy/min. Transplants were performed at least 5 hours after the last irradiation dose, with a variable number of Lin^- cells from a donor mouse. Lin^- donor cells were normally resuspended into a final volume of 200-300 µl of IMDM, and transplanted in the tail vein of the irradiated recipients. The effect of the irradiation upon the endogenous hematopoiesis was evaluated in control groups that were not infused with donor hematopoietic cells.

5.3. Engraftment analyses

PB collection from recipient mice was routinely carried out (every 30 days during three or four months post-transplant) by lateral tail vein sampling. A volume of 200 µl was collected. PB was mixed with 20 µl of EDTA (0.5M, pH=8) to prevent blood coagulation. At sacrifice, BM was also collected. Both femurs and tibias were surgically extracted, and BM was collected by flushing these bones with IMDM medium, under sterile conditions as previously mentioned in Materials & Methods section 3.2.3)

5.3.1. Chimerism by quantitative PCR (qPCR) in FVB FA-A mice

A syngeneic model of transplant was used to follow the hematopoietic repopulation of grafts from male donors in female recipients. This was evaluated by quantitative PCR (qPCR) in PB cells. The samples were lysed and the genomic DNA was extracted with the NucleoSpin® Tissue kit (Macherey-Nagel, Düren, Germany). Then, extracted DNA was quantified with a spectrophotometer (NanoDrop®) and adjusted to a concentration of 10 ng/µl.

Percentages of donor engraftment were calculated by reference to standard curves consisting of different percentages of male and female blood genomic DNA (gDNA) (from 0% to 100% male, 12 points with different percentage of donor were used) obtaining linear regression coefficients in the range of 0.95-0.99. Samples and standards were amplified in triplicate and analysed by qPCR. For the qPCR reaction, 4 µl of DNA (40 ng) were mixed with 10 µl of the Taqman Master Mix (Universal PCR Master Mix, Roche), 0.1 µM of the primers and 0.1 µM of the probes, specific for the male *Sry* gene or for mouse *β-Actin* gene. The conditions of the qPCR were: 95°C 10 minutes, 50 cycles of 95°C 30 seconds, and 58ºC 30 seconds. The qPCRs were performed in the Rotor-Gene 6 machine (Corbet, Australia). Primers and probes were generally generated and supported by Metabion (Grupo Taper, Munich, Germany) or by Sigma® Life Sciences (**Table 7**).

The percentage of engraftment of the male donor in female recipients was calculated with the 2(-ΔΔC(t)) Method) [313], a modified formula of the Livak method, assuming efficiency of 2 for both pair of primers, that allows the calculation of the expression of a gene relative to a housekeeping gene in a normalized manner. In our case, the housekeeping gene was the mouse *β-Actin*. The value obtained for each sample was extrapolated in the equation obtained from the data of the standard curve. Consequently, the percentage of male cells present in a female recipient was obtained. This percentage indicated the percentage of engraftment.

The formula used to calculate the engraftment is the one that follows, being Ct, the cycling threshold.

$$\%\ Donor\ Engraftment = 100\ x\ 2^{\Delta t} = 100\ x\ 2^{(Ct\ \beta Act - Ct\ Sry)}$$

5.3.2. Chimerism by flow cytometry in P3B mice

A congenic model of transplant was used to follow the hematopoietic repopulation of C57BL/6 ($CD45.2^+$) Lin^- donor cells in P3B ($CD45.1^+$) recipients. This was evaluated by flow cytometry in PB and BM samples. Cells were lysed and washed with PBA. Samples were stained with 0.5 µl of $CD45.1^+$ conjugated with PE (BD Pharmingen) and with 0.5 µl of the $CD45.2^+$ antibody conjugated with Biotin (BioLegend) (**Table 6**) for 30 minutes at 4ºC. Then, 0.8 µl/sample of SAV-PECy7 secondary antibody were added (Caltag), and samples were stained for 30 minutes at 4ºC. Once the samples were washed and resuspended in PBA, 1 µg/ml DAPI/PBA was added to the samples to evaluate the viability by flow cytometry. The hematopoietic engraftment was determined by flow cytometry measuring the percentage of $CD45.2^+$ cells against the endogenous reconstitution ($CD45.1^+$).

5.3.3. Study of the hematopoietic lineages in peripheral blood by flow cytometry

In the previous analysis of the chimerism by flow cytometry in P3B mice, PB samples were also analysed for multilineage engraftment. Myeloid cells were evaluated by staining the cells with 0.5 µl of Gr-1 and Mac-1 antibodies conjugated with PE; and lymphoid cells were evaluated by staining the cells with 0.5 µl of B220 and CD3-ε antibodies conjugated with PE-Cy5 (BD Pharmingen) (**Table 6**).

Purpose	Name	Sequence (from 5' to 3')	Tm (ºC)	Product size (pb)	Fluorochrome
qPCR *Sry*	Sry-F	TGTTCAGCCCTACAGCCACA	53.9	140	----------
	Sry-R	CCTCTCACCACGGGACCAC	54.8		----------
	Sry Probe	ACAATTGTCTAGAGAGAGCATGGAGGGCCA	64.7	----------	6-FAM
qPCR *β-Actin*	β-Actin-F	ACGGCCAGGTCATCACTATTG	53.9	131	----------
	β-Actin-R	ACTATGGCCTCAGGAGTTTTGTCA	55.9		----------
	β-Actin Probe	AACGAGCGGTTCCGATGCCCT	63.5	----------	Joe
PCR *T7E1* or *Cel1* assay	mAAVS1 CelIF	TCTGGATTCAGGATGCTTTT	54.0	405	----------
	mAAVS1 CelIR	TCACCTTGCTTCCACTTTCC	58.0		----------
5' PCR Integration junction	mAAVS1-5'F1	TTGTGGCCTCAGGACAGTGTAC	64.0	1167	----------
	mAAVS1-5'R1	AACGGACGTGAAGAATGTGCG	61.0		----------
	mAAVS1-5'F2	GGTTTCGTTCTCCTGCACTC	60.0	1439	----------
	mAAVS1-5'R2	GTCCGTCTGCGAGGGTACTA	63.0		----------
	mAAVS1-5'F3	TTCCCGTGACTTGTGCTGTA	64.4	1029	----------
	mAAVS1-5'R3	CCACGGGGTTGGGATTATTAT	65.2		----------
3' PCR Integration junction	mAAVS1-3'F1	ACAGATGGAAGGCCTCCTGG	63.0	1395	----------
	mAAVS1-3'R1	TCTTGGAACTTTCACTGCTAAAGC	62.0		----------
	mAAVS1-3'F2	GCAACCTCCCCTTCTACGAG	63.0	1337	----------
	mAAVS1-3'R2	GATGCCCCAAGGAGGGTTTA	60.0		----------
	mAAVS1-EGFP-3'F	GTGGTTTGTCCAAACTCATCAA	63.7	1035	----------
	mAAVS1-EGFP-3'R	TCCTTGTTTTCCTGGGACT	61.1		----------
PCR for sequencing 3' integration junction	mAAVS1-3'F2	GCAACCTCCCCTTCTACGAG	63.0	----------	----------
	mAAVS1-3'R2	GATGCCCCAAGGAGGGTTTA	60.0	----------	----------
	SequencingAAVS1-1_3'_F	TGGAGATGGGAAACAGAG	58.7	----------	----------
	SequencingAAVS1-2_3'_F	TGTCCCTAGAAGCTCTGGTG	62.0	----------	----------
	SequencingAAVS1-3_3'_F	CTGACTGCATCCCTCTCCTC	64.0	----------	----------
	SequencingAAVS1-4_3'_F	GCTTGGAAAGTGGAAGCAAG	63.8	----------	----------
	SequencingAAVS1-1_3'_R	TTGCATCTCCTTTCCCAATC	63.9	----------	----------

Table 7. Primers and probes used for different purposes in qPCRs and PCRs. Forward and reverse primers as well as probes, their sequence, their melting temperature (Tm) and the product size is indicated. In the case of probes, the fluorochrome to which each probe was conjugated is also indicated.

6. GENERAL GENE-EDITING PROTOCOL

6.1. Gene-editing protocol in cells lines and MEFs

For the gene editing of cell lines and MEFs these cells were cultured to achieve the required confluence or number of cells before transfection. In the case of HEK-293T cells, cells were lipofected, whereas the rest of cell types were nucleofected, and at 48 or 72 hours post-transfection cells were harvested for FACS analysis, for generating cell pellets for the *T7E1* or Surveyor assays or for other analyses such as western blot, WB. Analyses of targeted integration and phenotypic correction were only performed in immortalized FA-A MEFs. **Figure 11 A** summarizes the gene-editing protocol in HEK-293T cells, while **Figure 11 B** shows the gene-editing protocol followed in the rest of cell types (NIH/3T3, Ba/F3 and FA-A MEFs).

6.2. Gene-editing protocol in mouse HSPCs

Purified BM Lin^- cells were prestimulated for 48 hours in the corresponding medium at a density of $0.6x10^6$ cells/ml. Then, $1.4x10^6$ cells (unless otherwise stated) were nucleofected with the corresponding DNAs or mRNAs. At 24 or 48 hours post-transfection, cells were harvested, washed, counted and used for the different procedures: to analyse the viability and transfection efficiency by FACS (normally $5x10^4$ cells were used per condition), to generate cell pellets for Surveyor assay, or to establish CFCs cultures (with or without drugs for selection). For the *in vivo* analyses conducted at 48 hours post-transfection, most of the cells were used for transplant. In several experiments, cells were maintained in expansion for 14 days to conduct further analysis such as FACS analysis. **Figure 11 C** summarizes the gene-editing protocol in mHSPCs. The conditions of all gene-targeting experiments performed in mouse bone marrow hematopoietic progenitor cells (Lin^-) both from WT and FA-A FVB mice or from WT C57BL/6 mice are summarized in **Table S4** (Appendix 1).

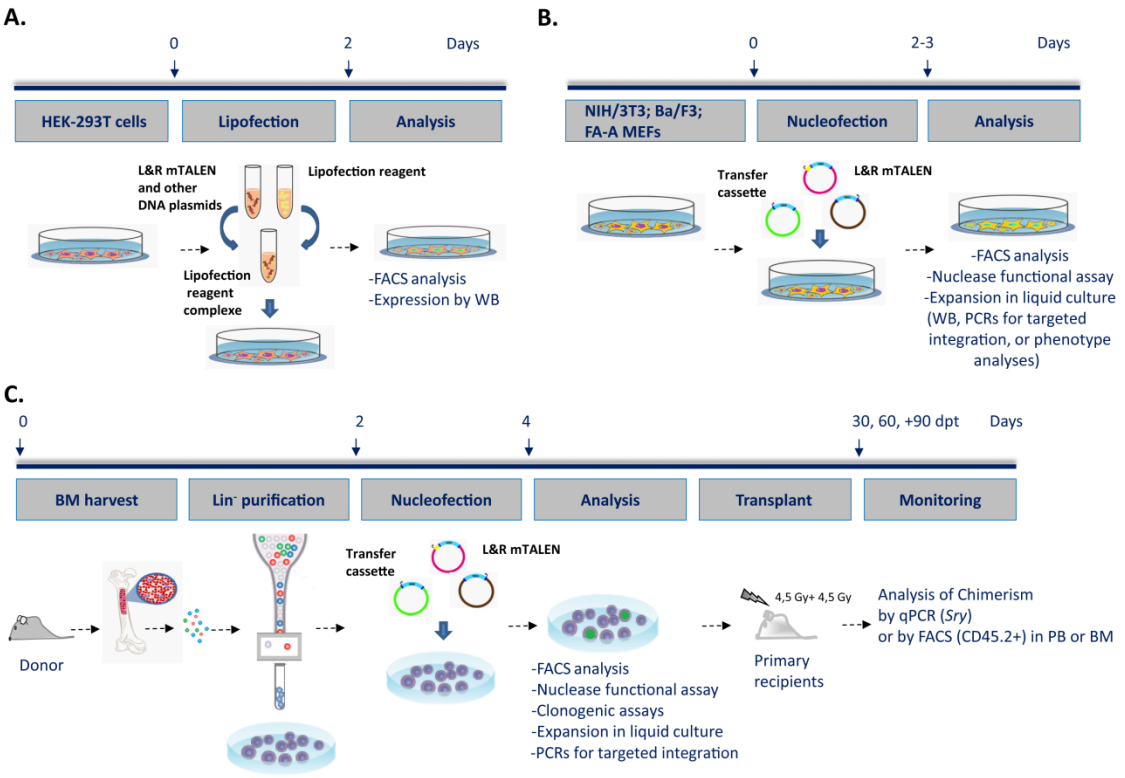


Figure 11. Scheme of the gene-editing protocol in different cell types. A) Protocol of gene editing in HEK-293T cells. B) Protocol of gene editing in NIH/3T3 cells, Ba/F3 cells and FA-A MEFs. C) Protocol of gene editing in mouse BM hematopoietic stem and progenitors cells, Lin$^-$ cells.

7. LIPOFECTION AND NUCLEOFECTION PROCEDURES

7.1. Transfection of HEK-293T with Lypofectamine

HEK-293T cells were seeded at 1x10^5 cells per well the day before lipofection. When HEK-293T cells reached 70% confluence they were lipofected with the L or R mTALEN monomers (of the NN- or NK-mTALEN pairs). Lipofectamine® 2000 Reagent was used according to the manufacturer's protocol (Invitrogen/ Life Technologies/ Thermo Fisher Scientific). Of each TALEN monomer 400 ng of DNA were co-transfected with 100 ng of an *EGFP* control plasmid and 500 ng of pUC118 control plasmid. At 6 hours post-transfection medium was changed with fresh medium. In all instances, plasmid DNA used for lipofections and nucleofections was endotoxin-free. At 48 hours post-nucleofection cells were harvested to conduct FACS analysis to determine the transfection efficiency (approximately 5x10^4 cells per condition were used for that purpose) and cell pellets were generated to conduct WB analyses.

7.2. Nucleofection

7.2.1. Nucleofection of NIH/3T3 cells

For nucleofection, 2x10^6 cells per condition were used with the Amaxa MEF2 Nucleofector® Kit (Lonza Group, Basel, Switzerland) using program T20 of the Nucleofector™ I device. Of each L and R monomer, 1.5 and 3 µg were nucleofected (of both the NK- and NN-mTALEN pairs). At 6 hours post-nucleofection the medium was changed with fresh medium. At

24 hours post-nucleofection cells were harvested to conduct FACS analysis to determine the transfection efficiency (approximately $5x10^4$ cells per condition were used for that purpose), and at 72 hours post-nucleofection, cell pellets were generated to conduct *T7E1* assay.

7.2.2. Nucleofection of Ba/F3 cells

For nucleofection, $2x10^6$ cells per condition were used with the Amaxa Cell Line Nucleofector® Kit V (Lonza Group) using program G-30 of the Nucleofector™ I device. Of each L and R monomer, 1 and 2 µg were nucleofected (of both the NK- and NN-mTALEN pairs). At 24 hours post-nucleofection 2/3 of the medium were replaced with fresh medium. At 72 hours post-nucleofection, the transfection efficiency was analysed by FACS (approximately $5x10^4$ cells per condition were used for that purpose) and cell pellets were generated to conduct the Surveyor assay.

7.2.3. Nucleofection of FA-A MEF cells

For nucleofection, $1x10^6$ cells per condition were used with the Amaxa Mouse/Rat Hepatocyte Nucleofector® Kit (Lonza Group) using program N-24 of the Nucleofector™ I device. This program had been validated by Lonza to work with primary (non-immortalized) mouse embryonic fibroblasts. Of each L and R monomer, 1.5 and 2.5 µg were nucleofected (of both the NK- and NN-mTALEN pairs). At 6 hours post-nucleofection the medium was changed with fresh medium. At 72 hours post-nucleofection, the transfection efficiency was determined by FACS (approximately $5x10^4$ cells per condition were used for that purpose) and cell pellets were generated to perform the *T7E1* assay.

7.2.4. Nucleofection of immortalized FA-A MEF cells

For nucleofection, $2x10^6$ cells per condition of immortalized FA-A MEFs were used, with the Amaxa MEF2 Nucleofector® Kit (Lonza Group) using program T20 of the Nucleofector™ I device. Of each mTALEN monomer, 2.5 µg, and different donor doses (0.75, 2 and 4 µg) were nucleofected. At 6 hours post-nucleofection the medium was changed with fresh medium. At 48 hours post-nucleofection, the transfection efficiency was determined by FACS, and cell pellets were generated to perform the Surveyor assay. Between $0.5x10^6$ and $0.8x10^6$ cells were maintained in expansion in treated 12-well/plates (22 mm, Corning) and maintained during 5 passages with puromycin (3 passages with 1 µg/ml and 2 passages with 1.25 µg/ml) and another 3 passages without puromycin. Cell pellets were then generated for target integration analyses and clones were generated by limiting dilution to perform further studies (described in the Materials & Methods section 9.1).

7.2.5. Nucleofection of Lin$^-$ BM cells

After purification, Lin$^-$ cells were prestimulated for 48 hours. Cells were seeded at a concentration of $6x10^5$ cells/ml in untreated 24-well/plates (16 mm) or in 6-well/plates (35 mm) (Corning, NY, USA) and cultured with the media and under the conditions indicated in **Table 5**. Then, cells were nucleofected with an *EGFP* control plasmid or the corresponding doses of the mTALEN monomers and donors.

In the two first experiments (Lin$^-$ 1 and Lin$^-$ 2), the Nucleofector Device I and large cuvettes were used for nucleofection, and $1\text{-}1.5x10^6$ cells were nucleofected per condition. The following programs were evaluated: X-01, U1, U8, U15, A13. The Nucleofection solution used was that from the Amaxa Mouse Macrophage Nucleofector® Kit (Lonza Group). In the

rest of experiments, the 4D-Nucleofector with large cuvettes was used, unless otherwise stated (in experiments Lin$^-$ 5 and Lin$^-$ 6 the strip cuvettes were used). In most of the experiments from 1x10^6 to 2x10^6 cells were nucleofected per condition. The P3 Primary Cell 4D-Nucleofector® X Kit (Lonza Group) was always used. Initially, different programs were evaluated: ED-113, DT-113, EO-100, DN-110, EA-105, ED-123, DS-123, EN-113, DS-118, and DK-100, but finally the ED-113 program was used in most nucleofections. Cells were left to recover from the nucleofection pulse in the cuvette with prewarmed culture medium containing serum and supplements but without antibiotics for 35-45 minutes in the incubator (depending on the experiment different cocktail media were used, see **Table S4,** Appendix 1). Next, cells were removed from the cuvette and plated at a concentration of 0.5-0.7x10^6 cells/ml in untreated 24 or 6-well/plates (16 mm or 35 mm, respectively) (Corning) and maintained in expansion in the incubator for 2 or 5 days depending on the experiment.

Nucleofection disorganized the lipid bilayer structure eventually, and not only the DNA or mRNA are introduced in the cells, but also other molecules that can lead to cell death [458, 528]. A molecule, Q-VD-OPh (also called QVD, Sigma® Life Sciences), that is a caspase inhibitor with potent antiapoptotic properties was added in some nucleofection experiments to avoid cell death and the toxicity associated to nucleofection [65]. When it was used, it was added 24 hours before nucleofection and just after the nucleofection at 50 µM (**Table S4**). Moreover, in one experiment (Lin$^-$ 39), two molecules proposed to increase the HRR events were added immediately after the nucleofection. We used the L755507 at 5 µM (Santa Cruz Biotechnology, Dallas, Texas, USA) and the BrefeldinA at 0.1µM (BioLegend). Doses were chosen based on published results in E14 mESCs [584]. Nucleofected cells were maintained in liquid culture with these molecules for 48 hours before the analysis of their effect.

8. NUCLEASE FUNCTIONAL ASSAY

The aim of this assay is to detect possible mismatches originated by the repair of DSBs generated by the cleavage of the TALEN in the *Mbs85* locus which, in the absence of a donor template, are preferentially repaired by NHEJ. When the DSBs are repaired by NHEJ, small insertions and deletions (also referred to as INDELS) are generated. A summary of the protocol followed in our studies is represented in **Figure 12**.

First, the gDNA of these cells was extracted using NucleoSpin® Tissue kit (Macherey-Nagel). Then, a PCR was performed to amplify the region of the *Mbs85* (mouse *AAVS1, mAAVS1*) locus where the mTALEN had cleaved. The primers used in this PCR were the mAAVS1 CelIF and mAAVS1 CelIR (primers and probes were generated and supported by Metabion (Grupo Taper, Munich, Germany) or by Sigma® Life Sciences. The list of primers is available in **Table 7**.

PCRs conducted for the *T7E1* assay were performed as follows: 200 ng of gDNA, 1 µl 10 µM of each primer, 1 µl 10 nM dNTPs, 10 µl of the Buffer 5X HF and 1 µl of Phusion® Pol enzyme (Phusion® High-Fidelity DNA Polymerase, New England Biolabs) in a final volume of 50 µl. The cycling conditions were the following: 98°C 30 seconds, 40 cycles of 98°C 20 seconds, 56°C 30 seconds and 72°C 30 seconds, and finally one cycle of 72°C 2 minutes. The majority of

the PCRs performed throughout this thesis have been done with the Veriti 96 Well Thermo Cycler machine (Applied Biosystems, Foster City, California, USA).

PCRs conducted for the Surveyor assay were performed as follows: 200 ng of gDNA, 1.25 µl 10 µM of each primer, 0.5 µl 100 nM dNTPs, 10 µl of the Buffer 10X and 1 µl of Herculase enzyme (Herculase II Fusion Enzyme, Agilent Technologies, Santa Clara, California, USA) in a final volume of 50 µl. The cycling conditions were the following: 95°C 2 minutes, 40 cycles of 95°C 20 seconds, 56°C 20 seconds and 72°C 30 seconds, and finally one cycle of 72°C 3 minutes.

PCR products were purified with the NucleoSpin® Gel and PCR Clean-up kit (Macherey Nagel). Both PCR products generated an amplicon of 405 bp that was then dehybridized and rehybridized in order to obtain heteroduplexes by denaturation and reannealing and formed between mutated and wild-type DNA strands. If an INDEL has occurred, the heteroduplex presents a DNA loop (hairpin) that can be recognized by the *T7* Endonuclease (*T7E1*) or the *Cel1* nuclease, generating a band pattern that can be visualized in gels. In our case, a positive result will be represented by the appearance of two additional bands, of 224 bp and 181 bp apart from the parental band of 405 bp.

When the *T7E1* endonuclease (New England Biolabs) was used, the whole purified PCR sample was heated up to 95°C for 5 minutes in a heat block, and left to cool down to RT (at least 2 hours) to produce the heteroduplexes. Then, 200 ng of the amplicon (in 11 µl of 1XNEB 2 buffer) and 1 µl of the *T7E1* (in a 1:1 mixture with NEB 2) were mixed up to a volume of 12 µl. Next, the reaction was incubated at 37°C for 20 minutes in a water bath, and the reaction was stopped by adding 1.2 µl 10X orange dye loading buffer and by putting the sample on ice. Next, the whole reaction was run on a 2% agarose gel until gaining a good resolution. Subsequently, it was analysed in a Molecular Imager® GelDoc™ XR$^+$ System (BioRad).

When the *Cel1* endonuclease (Surveyor® mutation detection kit, IDT, Coralville, Iowa, USA) was used, a PCR program with the purified PCR products was performed following the manufacturer's instructions to form heteroduplexes between mutant and wild-type strands. Then, 1 µl of Surveyor Nuclease *S* and 1 µl of Surveyor Enhancer *S* were added to the PCR product, and incubated for 45 min at 42°C. Thereafter, 1/10 Stop Solution was added to the reactions and loading buffer Novex Hi-Density TBE Sample Buffer (5X) was added to the samples (Invitrogen/ Life Technologies/ Thermo Fisher Scientific). The whole reaction was run on a 10% TBE polyacrylamide gel 1.0mm*10 well (Invitrogen/ Life Technologies/ Thermo Fisher Scientific) until gaining a good resolution. Finally, the gel was stained with Syber Gold Nucleid Acid Gel Stain (1:10000 TBE 1X) for 15-30 min and analysed in a Molecular Imager® GelDoc™ XR$^+$ System (BioRad).

The intensity of the images was analysed (in order to measure the percentage of cleavage) by measuring the densitometry value of the different bands, using the Quantity One Software (BioRad, Hercules, California, USA). The cleavage was quantified by calculating the ratio between the digestion products and the parental band, and then normalizing it with the result obtained in mock (transfected cells without DNA) or untransfected cells. The percentage of cleavage was determined by the next equation:

$$\% \text{ Cleavage (TALEN)} = \frac{\text{Digestion Products} - (2X \text{ Background})}{(\text{Parental Product} + \text{Digestion Products}) - (3X \text{ Background})} \times 100 - \% \text{ Cleavage (MOCK)}$$

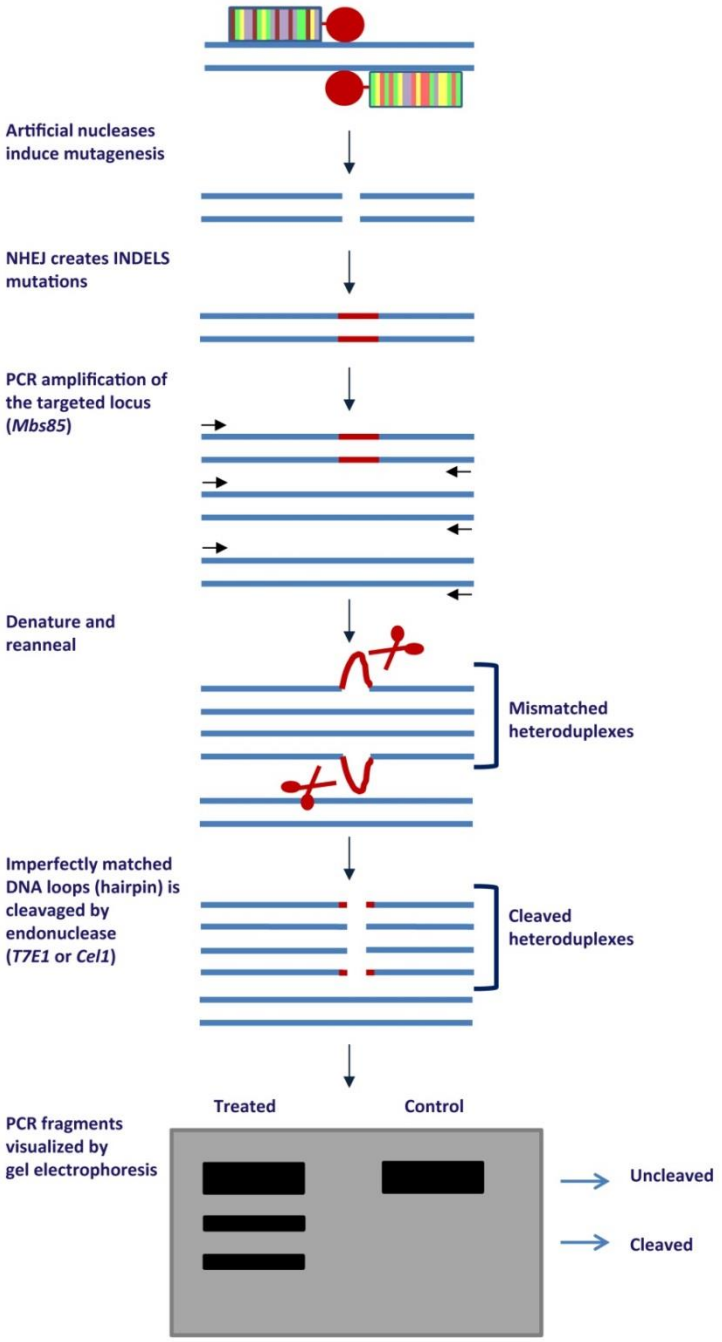


Figure 12. Summary of the experimental protocol of the *T7E1* or *Cel1* Mutation Detection assay.

9. ANALYSES OF THE ON-TARGET INTEGRATION

9.1. Generation of gene-edited FA-A MEF clones

From two different experiments, 174 clones were generated (in experiment 1, clones were derived from a bulk cell population nucleofected with 2.5 µg of each mTALEN monomer and 2 µg of donor template, and in experiment 2, clones were derived from a bulk cell population transfected with 2.5 µg of each mTALEN monomer and 0.75 µg of donor template). Clones were generated by limiting dilution, seeding 6 plates of 96-well (8 mm) per condition. Wells were seeded at 0.8 cells/well (8 cells/ml). Once the clones were established, they were kept growing in higher diameter plates, from 96-well to 24-well, 12-well and 6-well/plates.

9.2. PCR analyses for targeted integration of donors into the *Mbs85* locus

Genomic DNA from either the bulk cell population or from FA-A MEF clones was extracted using NucleoSpin® Tissue kit (Macherey-Nagel, Düren, Germany). Single colonies derived from CFCs assays (performed with Lin^- BM cells) were pelleted and resuspended in 10 µl of PBS. Their gDNA was extracted using 20 µl of lysis buffer (0.3 mM Tris HCl pH 7.5, 0.6 mM $CaCl_2$, 1,5 % Glycerol, 0.675 % Tween-20 and 0.3 mg/ml Proteinase K). Then, samples with the lysis buffer were incubated at 65°C for 30 min, 90°C for 10 min and next, maintained at 4°C. After lysis, 30 µl of water was added to the samples [90].

Two PCR reactions were conducted both for the 5' and the 3' integration junctions of the *Mbs85* integration site (**Figure 13**). In all instances, one primer hybridizes inside the donor, and another, into the *Mbs85* locus, outside of the HAs (in-out PCRs) (**Figure 13 A** and **B**). Different pairs of primers were designed to analyse the integration of the PGK-h*FANCA* donor: the pair mAAVS1-5'F (1 & 2) and mAAVS1-5'R (1 & 2); and the pair mAAVS1-3'F (1 & 2) and mAAVS1-3'R (1 & 2) (for more details of any of the primers used in these PCRs, see **Table 7**). Other primers were designed to analyse the integration of the PGK-*EGFP* donor: the pair mAAVS1-5'F3 and mAAVS1-5'R3, and the pair mAAVS1-EGFP-3'F and mAAVS1-EGFP-3'. Using this approach, amplification could only be detected when the donor template was integrated into the *Mbs85* locus.

PCR reactions were conducted using 200 ng of gDNA from bulk cell population or single MEF clones, or 15 µl of gDNA from single colonies. PCRs from the 5' and the 3' integration junctions were performed with 1.25 µl at 10 µM of each primer, 0.5 µl at 100 nM dNTPs, 10 µl of the Buffer 10X and 1 µl of Herculase enzyme (Herculase II Fusion Enzyme, Agilent Technologies) in a final volume of 50 µl. The cycling conditions were the following: 95°C for 10 minutes, 40 cycles of 95°C for 30 seconds, 59ºC -mAAVS1-5'F1&R1-; or 58ºC -mAAVS1-5'F2&R2-; or 62ºC -mAAVS1-5'F3&R3- for 60 seconds (for the 5' integration junction depending on the primers used); or 59ºC -mAAVS1-3'F1&R1-; 61ºC -mAAVS1-3'F2&R2-; 62ºC -mAAVS1-EGFP-3'F&R- for 60 seconds (for the 3' integration junction depending on the primers used); and 72°C for 1.5 minutes; and finally one cycle of 72°C for 10 minutes in a final volume of 50 µl.

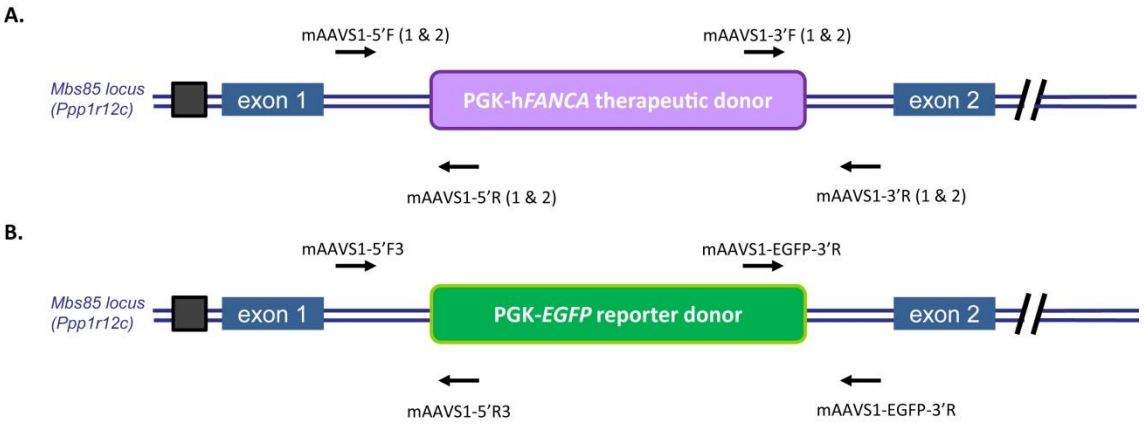


Figure 13. Representative scheme of the in-out PCRs to confirm the targeted integration into intron 1 of the *Mbs85* locus. PCRs and primers used to analyse the targeted integration (TI) in the 5' and the 3' integration junctions. The arrows indicate the forward and reverse pair of primers used for each integration junction. **A)** TI analysis of the PGK-h*FANCA* donor. **B)** TI analysis of the PGK-*EGFP* reporter donor.

9.3. Sequencing

PCR products of the 3' integration junction performed in the FA-A MEF bulk population and in the selected clones (#7#, #76#, #77#, #79# and #83#) were purified with the NucleoSpin® Gel and PCR Clean-up kit (Macherey Nagel) and different ratios were used to clone the PCR product into the pCR®-Blunt vector (Zero Blunt® PCR Cloning Kit, Invitrogen/Life Technologies/Thermo Fisher Scientific). The resulting product was transformed in TOP10 bacteria (Invitrogen, Life Technologies, Thermo Fisher Scientific) and plated in kanamycine plates. Plasmid DNA was purified from bacteria mini-cultures (generated from picked colonies) using the Fast Plasmid Mini Kit (5 Prime, Hilden, Germany) and was then analysed by digestions. Digestions with *EcoRI* were performed to analyse if the PCR products were correctly ligated, and MAXIpreps (NucleoBond® Xtra Midi EF, Macherey Nagel) were used for Sanger Sequencing (for each sequencing reaction, 15 µl of purified plasmid DNA at ≥100 ng/µl, and 5 µl/reaction of each 10 µM primer were necessary). Primers used for sequencing (sequenced in Stab Vida, Caparica, Portugal) were the followings ones: mAAVS1-3'-F2, mAAVS1-3'-R2; SequencingAAVS1-1_3'_F; SequencingAAVS1-2_3'_F; SequencingAAVS1-3_3'_F; SequencingAAVS1-4_3'_F; SequencingAAVS1-1_3'_R (for more details of any of the primers see **Table 7**). Analyses of the sequences were done with Finch TV version 1.4.0.

9.4. Fluorescence *in situ* hybridization (FISH)

Two sets of probes were designed and generated by Sandra Rodriguez-Perales (Molecular Cytogenetics Group, Centro Nacional Investigaciones Oncológicas, Madrid) who also performed the hybridization procedure. One probe recognized the target locus and another recognized the DNA PGK-h*FANCA* transgene. Since the *Mbs85* locus is located in chromosome 7 in the telomeric region A1, the probes RP23-336p21 and RP24-129K5 in bacterial artificial chromosomes (BACs) were designed to recognize the telomeric region E3 of chromosome 7, in order to prevent the recognition of the mTALEN target site.

To generate FISH probes, 1 µg of plasmid or BAC DNA was directly labelled using the Nick Translation Kit (Vysis, Abbott Molecular, Illinois, USA). This kit is designed for fluorescence labelling of DNA using fluorophore-labelled dUTPs (Spectrum Green or Spectrum Orange). The

labelled DNAs were co-precipitated with Cot-1DNA and DNA sheared salmon sperm (Vysis, Abbott Molecular) to prevent the unspecific hybridization in genomic repetitive DNA sequences and finally, the precipitated DNA mixture was resuspended in hybridization mix (Vysis, Abbott Molecular).

For FISH analysis, cells were treated with 100 µl/5 ml (0.02 µg/ml) of N-desacetyl-N-methylocolchicine (10 µg/ml) (KaryoMAX Colcemid Solution, Gibco/Thermo Fisher Scientific, Massachusetts, USA) to arrest cell cycle progression in metaphase state. Then, cells were treated with a hypotonic salt solution (KCl) and finally harvested after fixation in Carnoy solution (based on ethanol and glacial acetic acid). Next, cells were spread on a slide and maintained overnight at RT. The following day, samples were dehydrated through a series of ethanol washes and denatured in the presence of the probe/s on a plate at 75°C for 1 minute. Finally, samples and probes were hybridized overnight at 37°C. After hybridization washes with 0,4X SSC (Saline-Sodium Citrate buffer)/0,3% NP40 (2 min at 75ºC), 2X SSC/0,1% NP40 (5 min at RT) and PBD (Phosphate Buffered Detergent) (2 min at RT), chromosomes were counterstained with DAPI solution (Abbot Molecular) in Vectashield mounting medium (Vector Laboratories, USA).

10. WESTERN BLOT (WB) ANALYSES

Western blot (WB) analyses were performed to determine the expression of the HA epitope of mTALENs, and of the HPV 16 E7 protein in immortalized FA-A MEFs and gene-edited clones. Expression of the human FANCA protein was also analysed in immortalized FA-A MEFs and in gene-edited FA-A MEFs. Proteins were extracted from 5-7x10^6 cells per condition, using 200 µl RIPA buffer (150 mM NaCl, 50 mM TRIS, 1% NP-40, 0.1% SDS, 0.5% deoxycholate acid) with 4 µl of protease and phosphatase inhibitors (Mini-complete, Roche, Basel, Swiss) for HA detection. For HPV 16 E7 and hFANCA detection, proteins were extracted using 50 µl lysis buffer (2X) (40 mM Hepes (pH 7.5), 2% Triton, 200 mM NaCl, 40 mM $MgCl_2$, 20 mM EGTA, 80mM β-glycerophosphate and distillated water) and 50 µl protease and phosphatase inhibitors (2X) and 0.5 µl PMSF (200X) (1.74 gr with 50 ml Isopropanol). Finally, samples were centrifuged and the supernatants stored at -80ºC. Protein quantification was carried out using Lowry or Bradford methods (BioRad, Berkeley, California) in a spectrophotometer (Eppendorf Biophotometer AG 22331, Hamburg, Germany).

Of each sample, 40 or 50 µg were loaded in Bis-Tris 8% poliacrilamide gels (BioRad) or commercial NuPAGE 4-12% Bis-Tris poliacrilamyde gels (Novex/Life Technologies/Thermo Fisher Scientific) in redox and denaturing conditions with Laemmli buffer 1X [74, 285] or MOPS SDS running buffer 1X (NuPAGE, Novex/Life Technologies/Thermo Fisher Scientific) and electrophoresed (during different times for the different proteins at 100 V). Then, proteins were transferred to PDVF membrane for 2 hours at 90V and 4°C in 1X transfer buffer (25 mM Tris, 190 mM glycine, 20% methanol) or in commercial transfer buffer (NuPAGE, Novex/Life Technologies). After the transference, blocking of unspecific binding sites was performed with 5% powder milk diluted in PBS-T buffer (PBS 1X with 0.1% Tween-20) in agitation for 1 hour at RT. Then, the membrane was stained with the primary antibody diluted in 0.5% milk PBS-T with the following antibodies: anti-HA antibody produced in rabbit (1:1000, NB600-363, Novus

Biologicals, Littleton, Colorado, USA), anti-HPV16 E7 antibody produced in mouse (1:200, HPV16 E7 (NM2), sc-65711, Santa Cruz Biotechnology), anti-hFANCA antibody produced in rabbit (1:800, ab5063, Abcam, Cambridge, United Kingdom) or anti-β-ACTIN antibody produced in mouse (1:4000, ab6276, Abcam) overnight at 4°C. The membranes were washed with PBS-T buffer three times. Next, they were incubated with 1:5000 dilution in 0.5% milk PBS-T of the secondary antibody anti-rabbit (Horseradish Peroxidase (HRP)-conjugated anti-rabbit antibody, Dianova, Hamburg/ Germany; or pAb to Rb IgG-HRP, Abcam) or anti-mouse (Shp pAb to Ms IgG-HRT, Abcam) for 1 hour at RT. Finally, the membranes were washed three times with PBS-T. Detection was performed by Chemiluminiscence signal (SuperSignal West Pico Chemiluminescent Substrate, Thermo Fisher Scientific, Pierce, USA) or with the Clarity Western ECL Substrate (BioRad), with the ChemiDoc ™ MP Imaging System (BioRad) and processed with Image Lab ™ Software 5.2.1 (BioRad) or through photography film (Amersham, UK) exposure in an automatic developer (Curix60, AGFA, Mortsel, Belgium).

11. ASSAYS IN GENE-EDITED CELLS

These assays were performed in order to determine the functionality of the human *FANCA* gene in gene-edited FA-A MEFs.

11.1. MMC sensitivity curve

Gene-edited FA-A MEFs, both the bulk cell population and the selected clones, were exposed to MMC to investigate the potential correction of their phenotype. Per clone, 200 cells were seeded in treated 6-well/plates (35 mm) (Corning) in a final volume of 2 ml and increasing concentrations of MMC (0, 3, 10, 100, 300 and 1,000 nM). After one week, the medium was changed with fresh medium containing the same concentration of MMC. Cell viability was determined 14 days after cell seeding by scoring the number of colonies derived from 200 cells (CFUs, colony forming units). To facilitate the scoring of the colonies, plates were fixed with 4% paraformaldehyde (pH=7) (Panreac, Barcelona, España) (1.5 ml per each 6-well/plate) and after three washes with PBS 1X, wells were stained with hematoxylin (Thermo Fisher Scientific, Massachusetts, USA).

11.2. Chromosomal instability assay

Briefly, gene-edited FA-A MEFs, both the bulk cell population and the selected clones (#7#, #76#, #77#, #79# and #83#) were treated or not with 40 nM of MMC (Sigma® Life Sciences) for 24 hours to induce DNA damage. Then, 0.05 µg/ml of colcemid (KaryoMAX Colcemid Solution, Gibco/Thermo Fisher, Massachusetts, USA) was added to the culture medium and cells were incubated for 4 hours to arrest cell cycle progression in metaphase state. At that time, cells were treated with 0.56% of KCl (0.075M) to provoke an osmotic shock. The remaining analysis was done by Roser Pujol (Genome Instability and DNA Repair Group, Department of Genetics and Microbiology, Universitat Autònoma de Barcelona). Cells were fixed with methanol: glacial acetic acid (4:1) solution, centrifuged and Carnoy's solution was added to them in order to obtain an optimal concentration of fixed cells. Finally, cells were dropped onto clean slides. Extensions were made at 25ºC with 48% humidity in a Thermotron chamber, and the test of the concentration of cells was done by taking photographs with an inverted microscope. Metaphases were stained with 10% Giemsa stain, giemsa's azur eosin

methylene blue solution (Merck, New Jersey, USA) with Gibco® Gurr Buffer Tablets (Gibco/Invitrogen/Thermo Fisher Scientific). Then, cells were placed in glass coverslips with a pair of droplets of Entellan (Merck, New Jersey, USA). A minimum of 20 metaphases per culture were studied for the analysis of aberrations, although in some clones, a lower number of metaphases was studied (#7#, #77#, #79# and #83#).

12. IMAGE ANALYSES

In general, fresh culture cells were observed in an inverted microscope (Nikon Diaphot, Melville, NY, USA) and/or in an inverted fluorescence microscope (Olympus IX70 WH10X/22) to count and analyse $EGFP^+$ cells.

In FISH analyses, cell images were captured using a charge-coupled device (CCD) camera (Photometrics SenSys camera) connected to a computer with the Chromofluor image analysis system (Cytovision, Leica Biosystmes, UK).

Photographs taken of gene-edited FA-A MEFs of the MMC sensitivity curve assay were taken with a Sony Cyber Shot DSC F828 camera (Sony Corporation, Tokyo, Japan).

For the chromosomal instability assay, cell images were taken with an inverted microscope (Olympus CK30, USA).

All images were processed and organized using the Adobe Design Premium CS5 software.

13. OFF-TARGET ANALYSIS *IN SILICO*

Potential off-target sites for the NN-mTALEN pair were analysed with the PROGNOS software using TALENv2.0 algorithm (http://bao.rice.edu/cgi-bin/prognos/prognos.cgi). The 48 top off-target sites were ranked according to parameters optimized for TALEN off-target prediction. The mm10 genome was used in the search, and up to 6 mismatches were allowed in each nuclease half-site. The allowed spacing distances varied from 10 to 30 bp, and the formation of homodimers and heterodimers of the nucleases was included in the input.

14. STATISTICAL ANALYSES

Statistical analyses were performed using GraphPad Prism version 5.0 for Windows (GraphPad Software, San Diego, USA). Results are shown as the mean ± Standard Deviation (SD) or as the median ± interquartile range, from at least 3 replicates and from different experiments. The test of Kolmogorov-Smirnov was used to analyse the normality of the samples. When two sets of data were compared, the two-tailed t-Student or the Mann-Whitney tests were performed depending on whether or not the values followed a normal

distribution or not. When more than two sets of data were compared, the parametric one-way ANOVA followed by a post-hoc test (normally the multiple comparison Tukey test) or the non-parametric Kruskal-Wallis followed by the post-hoc Dunn's multiple comparison test were used. In some experiments, more than one variable was analysed. In these cases, when the data suit a normal distribution, the two-way ANOVA test was applied followed by a Bonferroni post-hoc test. Significances were indicated as P-value <0.05 (*), P-value <0.01 (**), P-value <0.001 (***) and P-value <0.0001 (****).

VIII. RESULTS

1. ANALYSIS OF EXPRESSION AND ACTIVITY OF DESIGNED TALEN THAT TARGET THE MOUSE *Mbs85* LOCUS

Two different pairs of TALE nucleases (TALEN) were designed and generated in the laboratory of Dr. Claudio Mussolino. These specific TALEN were designed to target the *Mbs85* (*Myosin Binding Subunit 85*) locus of the mouse genome, also known as *Ppp1r12c* (*Protein Phosphatase 1 Regulatory subunit 12C*). This locus is orthologous to the human *AAVS1 safe harbor* locus. These two pairs of mTALEN differ in the repeat-variable di-residues (RVD), in the targeting modules that recognize G nucleotides, although target the same sequence of the mouse *Mbs85* locus. Based on the amino acids of the targeting modules, mouse TALENs were referred to as the NN- and the NK-mTALEN pairs (described in Materials & Methods section 1.1).

1.1. Evaluation of mTALEN expression in HEK-293T cells

One day before lipofection, HEK-293T cells were seeded at $1x10^5$ cells per well. On the next day, when cells were at 70% confluence, samples were lipofected with both mTALEN monomers. The specific DNA plasmids carrying the left (L) and right (R) monomers that comprise each of the mTALEN dimers (the NN-TALEN pair and the NK-TALEN pair) were co-lipofected with an *EGFP* control plasmid in HEK-293T cells. On average, the viability was 59%, and the transfection efficiency obtained from the analysis of the EGFP expression was 36% at 48 hours post-transfection (**Figure 14 A**).

Cell pellets were then prepared to extract total protein lysates to confirm the functional expression of the different mTALEN monomers. Western blot analysis of the HA-tag (located in the extreme N-terminal of the backbone of the different mTALEN monomers) was performed. As shown in **Figure 14 B**, the expression of the HA-tag was detected in cells lipofected with the different monomers of the mTALEN.

This study demonstrates the efficient expression of the mTALEN constructs in HEK-293T cells.

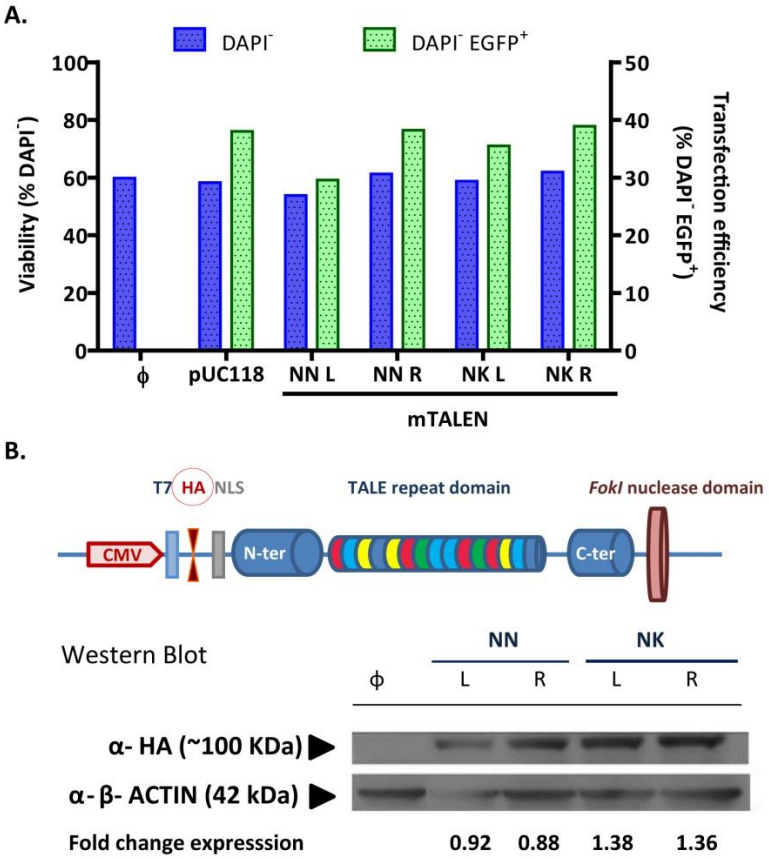


Figure 14. Evaluation of mTALEN expression in HEK-293T cells. **A)** Viability (measured as percentage of DAPI$^-$ cells) and transfection efficiency (percentage of EGFP$^+$ DAPI$^-$ cells) in the different conditions. ϕ: transfected cells without mTALEN plasmids (mock condition); pUC118: control vector; NN L, NN R, NK L, NK R: DNA plasmids carrying the different monomers, left (L) and right (R) monomers that comprise each of the TALEN dimers (the NN-TALEN pair and the NK-TALEN pair). **B)** Western Blot analysis of HA-tag present in each mTALEN monomer. The expression of the different monomers was calculated as a fold change with respect to β-ACTIN (used as a loading control).

1.2. Evaluation of mTALEN cleavage efficacy in different mouse cell types

To determine the activity of the different mTALEN pairs in different mouse cell lines and primary cells, genomic DNA was obtained from nucleofected cells with any of the mTALEN pairs. The cell types selected for these experiments were NIH/3T3 cells, Ba/F3 cells and FA-A MEFs (mouse embryonic fibroblasts from FA-A mice). Either the *T7E1* or the Surveyor assays were performed at 72 hours post-nucleofection to obtain a first estimate of the cleavage efficacy of the NN-mTALEN and the NK-mTALEN pairs that target the *Mbs85* locus.

The nucleofection procedures in NIH/3T3 cells ($2x10^6$ cells per condition), Ba/F3 cells ($2x10^6$ cells per condition) and FA-A MEFs ($1x10^6$ cells per condition) have been described in Materials & Methods section 7.2. Two different TALEN doses were tested in each cell type: from 3 to 6 µg of total DNA (L and R monomers) in NIH/3T3 cells, from 2 to 4 µg in Ba/F3 cells, and from 3 to 5 µg in FA-A MEFs. As a consequence of the *Cel1* or *T7E1* cleavage of the INDELs generated after NHEJ processing of DSBs in the parental locus (band of 405 bp), two different products of 224 and 181 pairs of bases (bp) were generated. As a control of the technique, samples were processed with and without the endonucleases *Cel1* or *T7E1*. Electrophoresis gels showed the efficiencies of INDEL generation among the different cell types with each

mTALEN dose. High percentages of cleavage were obtained in NIH/3T3 cells (between 12% and 39%), independently of the mTALEN pair. Although the NK-pair gave the highest percentage of cleavage (up to 39% cleavage activity) in these cells (**Figure 15 A**), the NN-pair also resulted in a high percentage of cleavage (up to 34% cleavage activity) using 3 µg of total DNA. In Ba/F3 cells and FA-A MEFs, the percentage of cleavage was always higher with the NN-mTALEN pair in comparison with the NK-pair; up to 15% versus a maximum of 6% in Ba/F3 (**Figure 15 B**), and up to 44% versus a maximum of 8% in FA-A MEFs (**Figure 15 C**). Based on the results obtained in FA-A MEFs with the NN-mTALEN pair at a dose of 2.5 µg per monomer, we chose this pair and this dose for the next set of nucleofection experiments.

These results thus demonstrate that mTALE nucleases efficiently cleave the *Mbs85* locus of different mouse cell types, including primary cells such as FA-A MEFs.

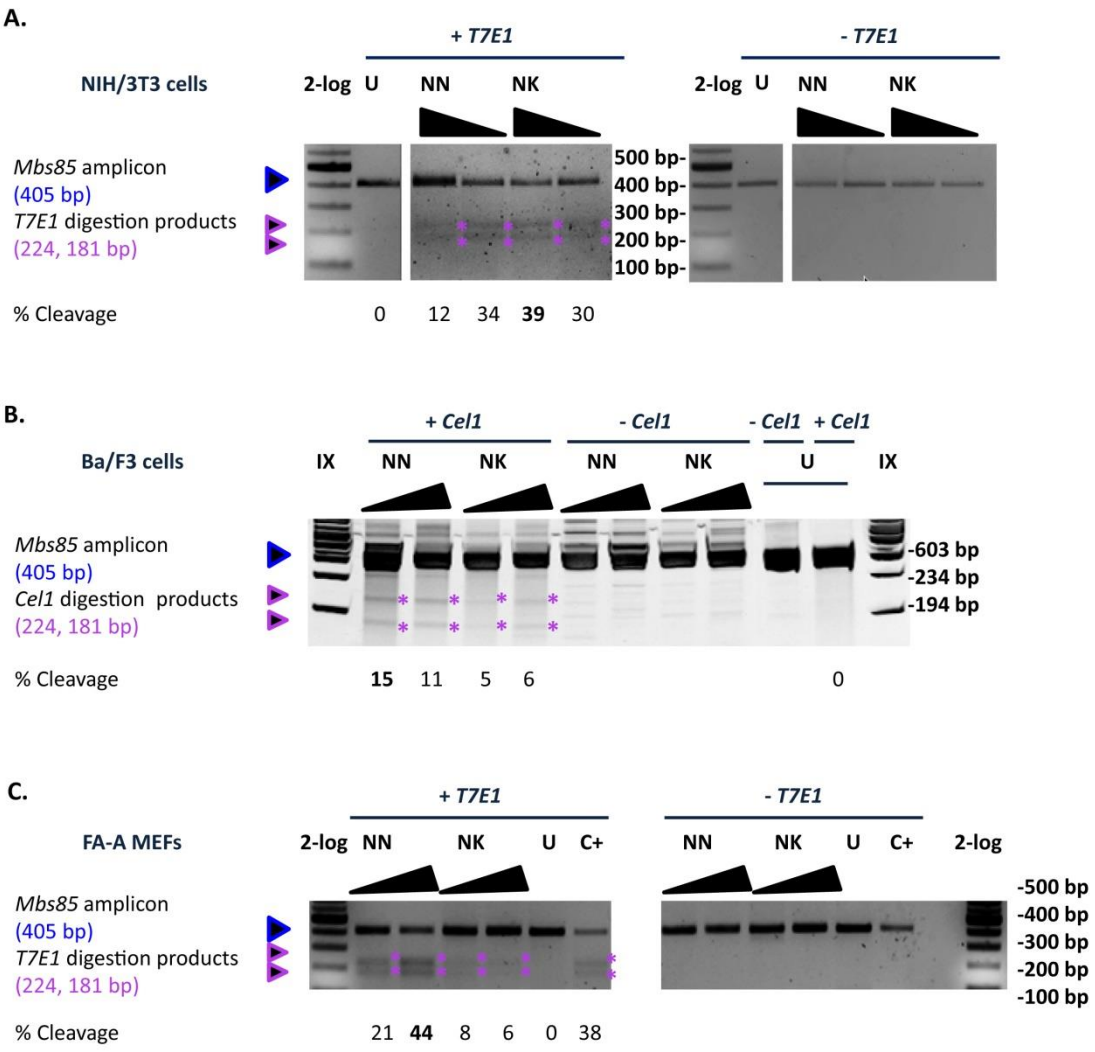


Figure 15. Evaluation of the cleavage efficacy of the indicated mTALEN pairs using the mismatch sensitive *T7E1* or Surveyor assays in NIH/3T3 cells, Ba/F3 cells and FA-A MEFs. Electrophoresis gels corresponding to **A)** NIH/3T3 cells, **B)** Ba/F3 cells and **C)** FA-A MEFs are shown in the figure. Disruption of the target locus was monitored using the mismatch sensitive *T7* Endonuclease I (*T7E1*) assay or the Surveyor assay (*Cel1*). Samples not digested with the endonuclease were used as controls. Two different concentrations of the mTALEN pairs were used (from 3 to 6 µg of total DNA in NIH/3T3 cells, from 2 to 4 µg of total DNA in Ba/F3 cells, and from 3 to 5 µg of total DNA in FA-A MEFs), and are indicated as a gradient in the figure. The extent of cleavage measured as the mean percentage of modified alleles is

indicated. Arrows indicate the size of the parental band (405 bp, in blue) and the expected positions of the digestion products (224 bp and 181 bp, in purple), that are also indicated with purple asterisks. The most active mTALEN pair is indicated in bold. C^+ is a positive control previously used in a *T7E1* or Surveyor assay. U: untransfected condition; IX and 2-log: DNA molecular weight markers.

1.3. *In silico* analysis to determine possible off-targets of the NN mTALEN pair

The NN-mTALEN pair was designed specifically to target the mouse *Mbs85* locus. However, we were interested in determining *in silico* the potential off-targets of the NN-mTALEN pair.

With that purpose, we used the recently developed PROGNOS software with the TALENv2.0 algorithm (http://bao.rice.edu/cgi-bin/prognos/prognos.cgi). When the bioinformatics tool was run, 48 top off-target sites were identified for the *Mbs85*-specific mTALEN in the mm10 mouse genome (although up to 4,893 potential mismatches were located). As was recommended, in the input we considered up to 6 mismatches in each nuclease half-site (monomer). Additionally, we allowed spacing distances of between 10 and 30 bp and we included homodimers and heterodimers formed by the nucleases. **Table 8** summarizes the 4,893 potential mismatches types that could be generated with both monomers (left and right) considering the different combinations of mismatches in each monomer, and the genomic regions where these off-targets are located. **Table 9** summarizes the possible 48 top off-targets identified for the NN-mTALEN pair with up to 6 mismatches in each nuclease half-site.

In subsequent experiments the specificity of our *Mbs85*-specific NN-mTALEN designed nucleases will be investigated by high-throughput sequencing of the top 20-25 off-targets.

4,893 total potential off-targets	
Mismatch types for each mTALEN monomer and number of potential off-target sites	**Genomic Regions**
0 and 0: 1 3 and 5: 2 4 and 4: 3 3 and 6: 10 4 and 5: 21 4 and 6: 170 5 and 5: 81 5 and 6: 1,089 6 and 6: 3,516	Exons: 136 Promoters: 26 Introns: 1,638 Intergenic: 3,093

Table 8. Summary of the mismatches types and the genomic regions of the off-targets obtained *in silico* with the NN-mTALEN pair.

Ranking	Nucleotide differences			Closest Gene	Region	Chromosome	Match Type	Left Target (5'-->3`)	Spacer (bp)	Right Target (3'-->5`)
	Total	Left	Right							
1	0	0	0	*Ppp1r12c*	Intron	chr 7	R13L	TCTGTTCCCACCAGTAACT	13	GATCGTTCTTCTCTCCTGT
2	9	3	6	*Vstm2b*	Intron	chr 7	R20L	TCTGTTCCCACCAGTAACT	20	GATCGTTCTTCTCTCCTGT
3	8	4	4	*Gcnt4*	Intergenic	chr 13	L16R	TGTCCTCTCTTCTTGCTAG	16	TCAATGACCACCCTTGTCT
4	8	3	5	*Opcml*	Intron	chr 9	L27L	TGTCCTCTCTTCTTGCTAG	27	GATCGTTCTTCTCTCCTGT
5	8	4	4	*Ush2a*	Intron	chr1	L15R	TGTCCTCTCTTCTTGCTAG	15	TCAATGACCACCCTTGTCT
6	9	6	3	*Hs3st3b1*	Intergenic	chr 11	L24R	TGTCCTCTCTTCTTGCTAG	24	TCAATGACCACCCTTGTCT
7	9	6	3	*Gm15800*	Intron	chr 5	R27L	TCTGTTCCCACCAGTAACT	27	GATCGTTCTTCTCTCCTGT
8	9	5	4	*Pga5*	Promoter	chr 19	R26R	TCTGTTCCCACCAGTAAC	26	TCAATGACCACCCTTGTCT
9	10	6	4	*4933402J15Rik*	Intergenic	chr 14	L22R	TGTCCTCTCTTCTTGCTAG	22	TCAATGACCACCCTTGTCT
10	9	6	3	*Odz4*	Intergenic	chr 7	L25R	TGTCCTCTCTTCTTGCTAG	25	TCAATGACCACCCTTGTCT
11	10	4	6	*Acbd6*	Intron	chr 1	L27L	TGTCCTCTCTTCTTGCTAG	27	GATCGTTCTTCTCTCCTGT
12	10	4	5	*Nkx6-1*	Intergenic	chr 5	R22L	TCTGTTCCCACCAGTAACT	22	GATCGTTCTTCTCTCCTGT
13	10	6	4	*Adam32*	Intron	chr 8	R15L	TCTGTTCCCACCAGTAACT	15	GATCGTTCTTCTCTCCTGT
14	9	6	3	*Cdh11*	Intergenic	chr 8	R10R	TCTGTTCCCACCAGTAACT	10	TCAATGACCACCCTTGTCT
15	9	5	4	*Gm5072*	Intergenic	chr X	L21R	TGTCCTCTCTTCTTGCTAG	21	TCAATGACCACCCTTGTCT
16	9	4	5	*Cry1*	Intergenic	chr 10	R29L	TCTGTTCCCACCAGTAACT	29	GATCGTTCTTCTCTCCTGT
17	9	4	5	*Il20ra*	Intron	chr 10	L12L	TGTCCTCTCTTCTTGCTAG	12	GATCGTTCTTCTCTCCTGT
18	9	5	4	*Sec11c*	Intergenic	chr 18	R18R	TCTGTTCCCACCAGTAACT	18	TCAATGACCACCCTTGTCT
19	10	6	4	*Tmem132b*	Intergenic	chr 5	L22R	TGTCCTCTCTTCTTGCTAG	22	TCAATGACCACCCTTGTCT
20	8	4	4	*Atp6v1h*	Intergenic	chr 1	L28R	TGTCCTCTCTTCTTGCTAG	28	TCAATGACCACCCTTGTCT
21	10	6	4	*Rab2a*	Intergenic	chr 4	R12R	TCTGTTCCCACCAGTAACT	12	TCAATGACCACCCTTGTCT
22	9	3	6	*Luc7l*	Exon	chr 17	R30L	TCTGTTCCCACCAGTAACT	30	GATCGTTCTTCTCTCCTGT
23	10	4	6	*Akr1b7*	Intron	chr 6	L19L	TGTCCTCTCTTCTTGCTAG	19	GATCGTTCTTCTCTCCTGT
24	10	6	4	*Pvt1*	Intergenic	chr 15	L12R	TGTCCTCTCTTCTTGCTAG	12	TCAATGACCACCCTTGTCT

25	10	6	4	*Fkbp2*	Intron	chr 19	L21L	TGTCCTCTCTTCTTGCTAG	21	GATCGTTCTTCTCTCCTGT
26	10	4	6	*Pax8*	Intron	chr 2	R15R	TCTGTTCCCACCAGTAACT	15	TCAATGACCACCCTTGTCT
27	10	6	4	*Gm11426*	Intergenic	chr 11	L12R	TGTCCTCTCTTCTTGCTAG	12	TCAATGACCACCCTTGTCT
28	9	5	4	*Izumo3*	Intergenic	chr 4	L20R	TGTCCTCTCTTCTTGCTAG	20	TCAATGACCACCCTTGTCT
29	9	4	5	*Tle3*	Intergenic	chr 9	L12R	TGTCCTCTCTTCTTGCTAG	12	TCAATGACCACCCTTGTCT
30	9	4	5	*Dsel*	Intergenic	chr 1	R27R	TCTGTTCCCACCAGTAACT	27	TCAATGACCACCCTTGTCT
31	10	4	6	*Sec23a*	Intergenic	chr 12	R18L	TCTGTTCCCACCAGTAACT	18	GATCGTTCTTCTCTCCTGT
32	11	6	5	*Gabra3*	Intron	chr X	L25R	TGTCCTCTCTTCTTGCTAG	25	TCAATGACCACCCTTGTCT
33	10	5	5	*Gabra6*	Intergenic	chr 11	R30L	TCTGTTCCCACCAGTAACT	30	GATCGTTCTTCTCTCCTGT
34	9	5	4	*Cntnap3*	Intergenic	chr 13	L17L	TGTCCTCTCTTCTTGCTAG	17	GATCGTTCTTCTCTCCTGT
35	9	5	4	*Gab3*	Intron	chr X	L12L	TGTCCTCTCTTCTTGCTAG	12	GATCGTTCTTCTCTCCTGT
36	10	5	5	*5730469M10Rik*	Intergenic	chr 14	L30L	TGTCCTCTCTTCTTGCTAG	30	GATCGTTCTTCTCTCCTGT
37	10	4	6	*Ctnnd2*	Intergenic	chr 15	L14L	TGTCCTCTCTTCTTGCTAG	14	GATCGTTCTTCTCTCCTGT
38	10	6	4	*Ggta1*	Intron	chr 2	L24L	TGTCCTCTCTTCTTGCTAG	24	GATCGTTCTTCTCTCCTGT
39	10	5	5	*4921511C10Rik*	Intergenic	chr3	R12R	TCTGTTCCCACCAGTAACT	12	TCAATGACCACCCTTGTCT
40	10	4	6	*4933430M04Rik*	Intergenic	chr 11	R25L	TCTGTTCCCACCAGTAACT	25	GATCGTTCTTCTCTCCTGT
41	9	6	3	*Luzp2*	Intron	chr 7	L11R	TGTCCTCTCTTCTTGCTAG	11	TCAATGACCACCCTTGTCT
42	11	6	5	*Efcab7*	Intron	chr 4	L27L	TGTCCTCTCTTCTTGCTAG	27	GATCGTTCTTCTCTCCTGT
43	9	6	3	*Rspo3*	Intergenic	chr 10	L26R	TGTCCTCTCTTCTTGCTAG	26	TCAATGACCACCCTTGTCT
44	9	4	5	*Hpcal1*	Intergenic	chr 12	L21L	TGTCCTCTCTTCTTGCTAG	21	GATCGTTCTTCTCTCCTGT
45	11	6	5	*Myo1b*	Intergenic	chr 1	R25L	TCTGTTCCCACCAGTAACT	25	GATCGTTCTTCTCTCCTGT
46	10	5	5	*Vps41*	Intron	chr 13	R26L	TCTGTTCCCACCAGTAACT	26	GATCGTTCTTCTCTCCTGT
47	10	4	6	*Atp10a*	Intergenic	chr 7	L16R	TGTCCTCTCTTCTTGCTAG	16	TCAATGACCACCCTTGTCT
48	10	6	4	*Adh6-ps1*	Intergenic	chr 3	L24L	TGTCCTCTCTTCTTGCTAG	24	GATCGTTCTTCTCTCCTGT

Table 9. Summary of the potential 48 top off-targets for the NN-mTALEN pair considering up to 6 mismatches in each nuclease half-site. The top off-targets are listed according to the ranking generated with TALEN v.2 algorithm. The nucleotide differences in the left and in the right locations, the closest gene, the region and the chromosome of the off-target localization are indicated. The match type and the sequence of the left and the right target are also included. Mismatches with the *Mbs85* (*Ppp1r12c*) locus are highlighted in red letters.

2. GENE-TARGETING STUDIES IN FA-A MEFs

2.1. Immortalization of FA-A MEFs

FA-A MEFs frequently develop cellular senescence after a few passages in culture. Therefore, we immortalized FA-A MEFs in order to generate a population that could avoid normal cellular senescence and hence, that could be maintained in culture during our gene-editing studies.

Immortalization of WT and FA-A MEFs was achieved after two consecutive transductions with a gamma-retroviral vector encoding the papillomavirus *16 E6-E7* genes. This procedure has been described in Materials & Methods section 3.2.2. Western blot analysis of the HPV 16 E7 protein confirmed the E7 expression in bulk cell populations of WT and FA-A MEFs and also in FA-A gene-edited clones generated in experiments described in the following sections (**Figure 16 A**).

To corroborate that MEFs immortalization did not change the characteristic phenotype of FA cells, in particular their hypersensitivity to interstrand cross-linking agents such as MMC, the resistance to MMC was tested both in WT and FA-A immortalized MEFs. Cells were seeded at 200 cells per well in 6-well/plates. The following day, increasing doses of MMC were added to the cells, which were maintained for 14 days in culture. The number of colonies was scored as described in Material and Methods section 11.1. The survival curves shown in **Figure 16 B** indicate that immortalized FA-A MEFs preserved the characteristic hypersensitivity to MMC of FA-A cells.

Thus, differences in the MMC sensitivity between WT and FA-A MEFs were maintained despite the immortalization process of the respective cell types.

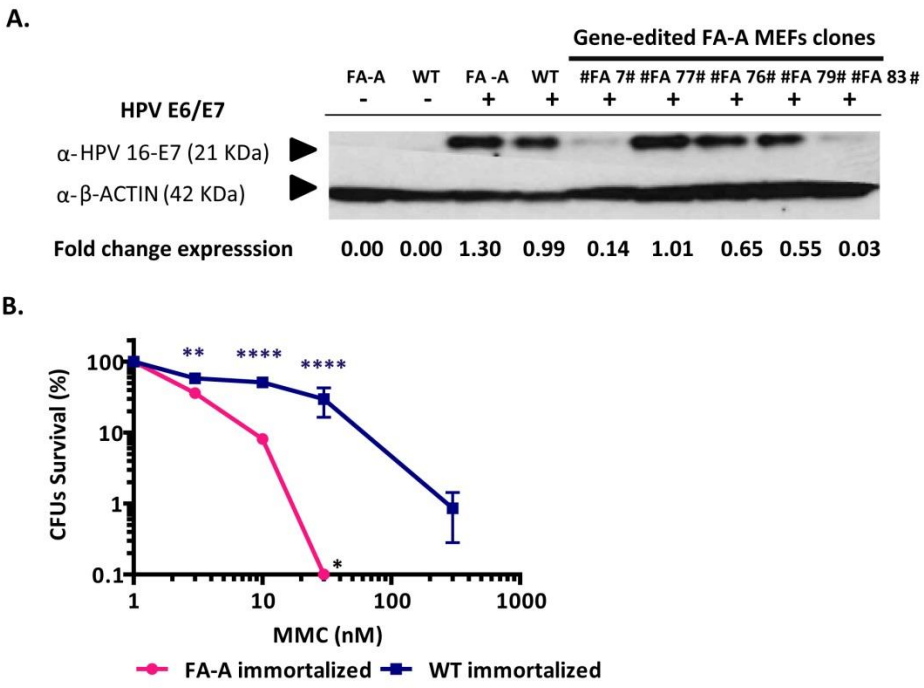


Figure 16. Immortalization of WT and FA-A MEFs and analysis of their sensitivity to MMC. **A)** Western blot analysis of HPV16 E7 protein. Analysis performed in immortalized FA-A and WT MEFs, and in gene-edited FA-A MEF clones (#7#, #76#, #77#, #79# and #83#). The expression was calculated as a fold change with respect to β-ACTIN, used as a loading control. **B)** Survival curve of FA-A and WT immortalized cells. Black asterisk (*) indicates that no colonies were generated in this condition and shows the highest survival that would be observed at 30 nM of MMC corresponding to the growth of a single colony in FA-A immortalized cells. (**) P-value <0.01, and (****) P-value <0.0001 indicate significant differences between FA-A and WT MEFs with a two-way ANOVA test followed by a post-hoc Bonferroni test.

2.2. Evaluation of the viability, transfection efficiency and NN-mTALEN cleavage efficacy in FA-A MEFs nucleofected with mTALEN and donor plasmids

A DNA donor carrying the h*FANCA* gene, the PGK-h*FANCA*-2A-*PuroR*-SV40pA, referred to as the PGK-h*FANCA* donor, was nucleofected together with the NN-mTALEN pair to evaluate targeted integration events into the *Mbs85* target site of FA-A MEFs.

Initially, we estimated the efficiency of nucleofection of the NN-mTALEN pair and the therapeutic PGK-h*FANCA* donor. With this purpose, one million cells were nucleofected per condition with 2.5 µg of each mTALEN monomer and different donor doses (0.75, 2 and 4 µg). Due to the lack of any reporter gene in the donor or in the NN-mTALEN constructs, an *EGFP* control plasmid was nucleofected to evaluate the transfection efficiency. In all instances, analyses were performed at 48 hours post-nucleofection. The average transfection efficiency was 30% (obtained in cells nucleofected with the aforementioned *EGFP* plasmid). The viabilities ranged from 47% to 78%. The use of higher doses of the donor was associated with lower viabilities, although no significant differences were obtained among the different conditions (**Figure 17 A**).

To test the cleavage efficacy of the NN-mTALEN in immortalized FA-A MEFs, Surveyor analyses were conducted. On average, 22.61 ± 11.55% of INDELS (data from five independent experiments) were generated in the *Mbs85* locus when only the mTALEN pair was used (see representative experiment in **Figure 17 B,** condition T). To evaluate whether the generation of INDELs was modified in the presence of the donor, a Surveyor assay was also conducted in cells that were co-transfected both with the mTALEN and the donor. Our results showed a reduction in the percentage of INDELS recognized by the *Cel1* endonuclease when the donor construct was also present (**Figure 17 B,** conditions T+D).

Taken together, our results show that nucleofection of the NN-mTALEN pair generated high percentages of cleavage (between 10% and 36%) in the *Mbs85* locus of FA-A MEFs.

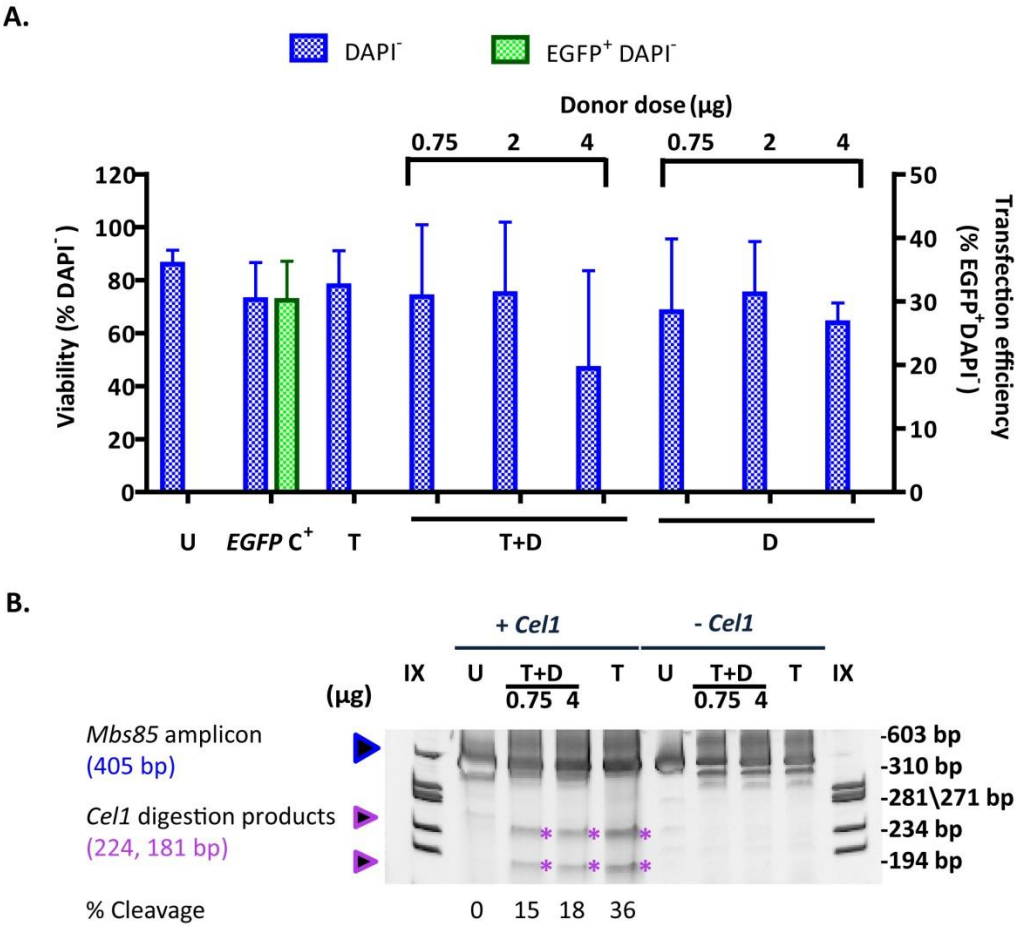


Figure 17. Evaluation of the viability and transfection efficiency of nucleofected FA-A MEFs, and analysis of the NN-mTALEN cleavage efficacy in the *Mbs85* locus. A) Analysis of viability (percentage of DAPI⁻ cells) and transfection efficiency (DAPI⁻ EGFP⁺ cells) in the different conditions. U: untransfected cells; *EGFP* C⁺: *EGFP* control plasmid; T: 2.5 µg of each NN-mTALE nuclease monomer; T+D: 2.5 µg of each NN-mTALE nuclease monomer plus therapeutic PGK-hFANCA donor doses indicated (0.75, 2 or 4 µg); D: different donor doses (0.75, 2 or 4 µg). Bars indicate the mean ± S.D. (n=2 experiments). B) Representative electrophoresis gel showing the disruption of the target locus in FA-A MEFS nucleofected with only the mTALEN (T, 2.5 µg of each NN-mTALE nuclease monomer) or together with different donor doses (0.75 µg and 4 µg). Samples not digested with the endonuclease were used as controls. The extent of cleavage, measured as the mean percentage of modified alleles, is indicated. Arrows indicate the size of the parental band (405 bp, in blue) and the expected positions of the digestion products (224 bp and 181 bp, in purple), that are also indicated with purple asterisks. IX: DNA molecular weight marker.

2.3. Targeted integration efficiency of the therapeutic PGK-h*FANCA* donor into the *Mbs85* locus of FA-A MEFs

FA-A MEFs that had been nucleofected either with the NN-mTALEN pair and the donor or with only the donor were maintained in culture for 5 passages in the presence of puromycin (1-1.25 µg/ml). This selection was used to enrich the population in gene-edited cells, since the donor cassette contained a *puroR* gene. Our data suggested that the puromycin selection was not very efficient in our system, as several clones generated after puromycin selection resulted negative for targeted integration. To evaluate the integration efficiency of the therapeutic PGK-h*FANCA* donor into the *Mbs85* locus, several PCRs were performed with the genomic DNA extracted from these cells. Two different PCRs were designed to analyse the 5' and the 3' integration junctions. For these analyses, we used primers hybridizing inside the donor

construct and outside the left and right HAs (in-out PCRs). **Figure 18 A** shows the scheme of the designed PCRs and **Figure 18 B** is a representative PCR from experiment 2 showing the amplicons of the in-out PCRs performed in FA-A MEFs. Interestingly, amplification of the specific bands corresponding to the insertion of the donor into the *Mbs85* site was only obtained in samples that had been treated both with the therapeutic donor and the mTALEN (except in experiment 2, in one case where the mTALEN were cotransfected with 2 µg of donor and where we could not confirm that the band corresponded to the integration band). No bands were observed in samples nucleofected with only the donor.

Aiming to determine the efficiency of targeted integration in FA-A MEFs, a total of 174 clones were generated by limiting dilution without adding puromycin to the culture medium in two independent experiments (explained in Materials & Methods section 9.1). In experiment 1, clones were derived from FA-A MEFs that had been nucleofected with 2.5 µg of each mTALEN monomer and 2 µg of donor template. In experiment 2, a lower dose of donor template was used to nucleofect FA-A MEFs, 0.75 µg. The integration efficiency of the therapeutic PGK-h*FANCA* cassette in each clone was analysed both for the 5' and the 3' integration junctions. The efficiency of TI obtained from the analyses of the 3' integration junction was 27.45% and 27.90% in experiments 1 and 2, respectively, whereas the TI efficiency corresponding to the 5' integration junction was 1.96% and 7.35% in these experiments. Because all clones that were positive for the 5' integration junction were also positive for the 3' integration junction, the efficiency of TI calculated for both integration junctions was 1.96% and 7.35% in experiments 1 and 2, respectively. A summary of the estimated HDR frequencies corresponding to these experiments is shown in **Table 10.** Five clones derived from experiment 2 (#7#, #76#, #77#, #79# and #83# clones) were selected for further analyses in the next set of experiments. **Figure 18 C** shows a representative PCR analysis of the 3' integration junction with four of these gene-edited FA-A MEF clones.

To confirm that the integration of the donor into the *Mbs85* site was the correct one, the band that appeared in the 3' integration junction PCR was cloned and analysed by Sanger sequencing (**Figure 18 D**). The nucleotide alignment was done by BLAST. The sequence obtained with one of the primers used for the sequencing of the 3' integration junction of clone #76# is represented in **Figure S1** from Appendix 1. The expected sequence of the 3' integration junction was also confirmed by sequencing of the 3' integration junction PCR in the bulk population of FA-A MEFs.

These experiments demonstrate the targeted integration of the therapeutic h*FANCA* donor into the *Mbs85* locus of FA-A MEFs. The overall efficiency of targeted integration varied between 1.96% and 7.35%.

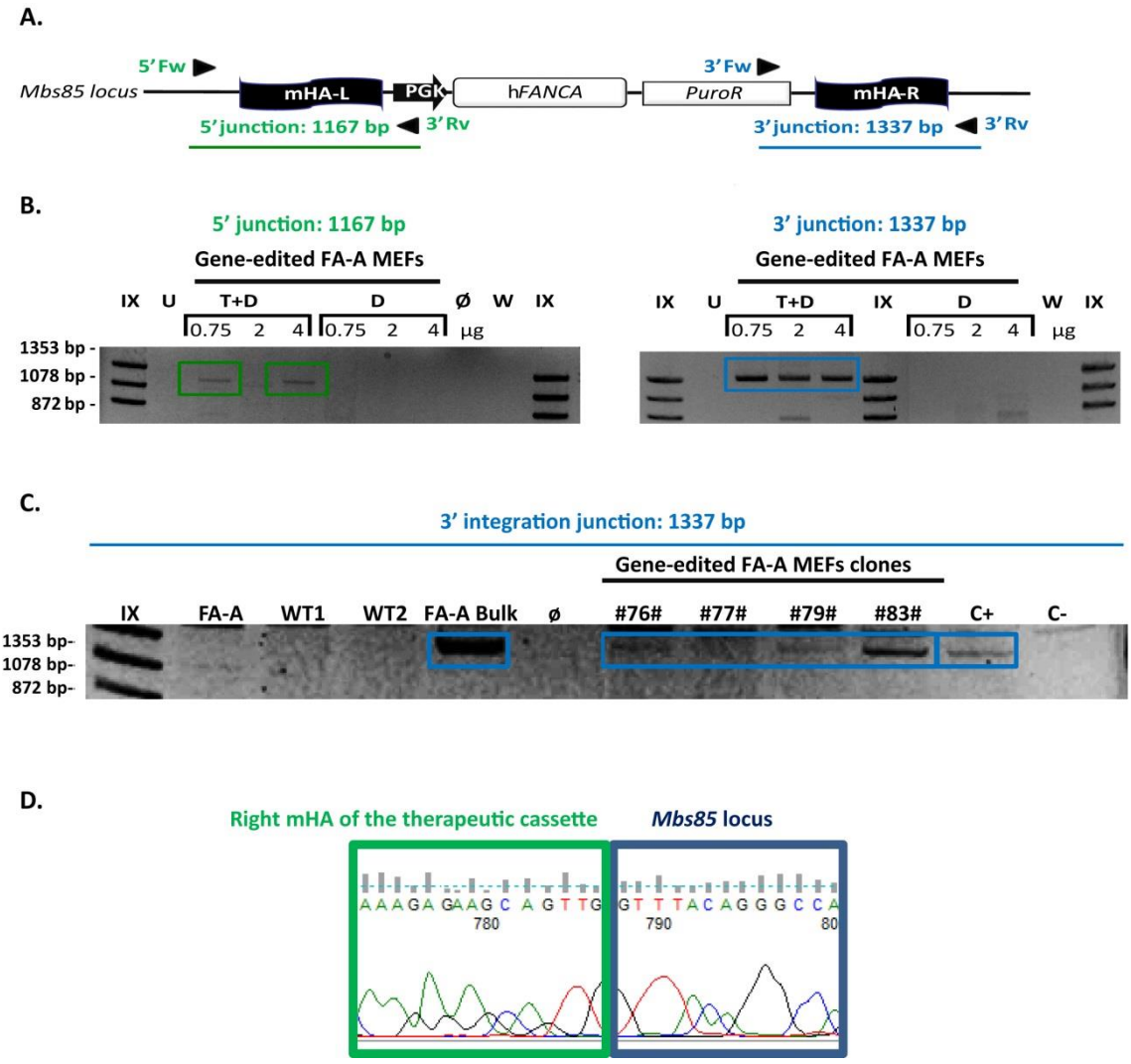


Figure 18. Targeted integration analyses in FA-A MEFs. **A)** Schematic representation of the targeted integration of the therapeutic PGK-*hFANCA* donor into the *Mbs85* locus of FA-A MEFs. Arrows represent the primers, forward (Fw) and reverse (Rv) used to evaluate the site-specific integration. The size of the PCR product is indicated for each integration junction, the 5' and the 3'. **B)** Electrophoresis gel representing a PCR analysis of experiment 2, from left to the right, the 5' and the 3' integration junctions. U: untransfected cells; ø: no sample; W: water control of PCR; T: 2.5 µg of each NN-mTALE nuclease monomer; T+D: 2.5 µg of each NN-mTALE nuclease monomer plus the donor doses indicated (0.75, 2 or 4 µg); D: 0.75, 2 or 4 µg of donor dose; and IX: DNA molecular weight marker. **C)** Representative 3' integration junction PCR of the bulk population of gene-edited FA-A MEFs (T+D0.75 µg) and in four clones: #76#, #77#, #79# and #83#. **D)** Sequenced region of clone #76# displaying part of the chromatogram that represents the designed right mHA (green) of the therapeutic PGK-hFANCA cassette, and part of the right location of the *Mbs85* locus (blue) (3' integration junction).

Experiment 1	5’	3’	Experiment 2	5'	3’
Total number of clones	54	54	Total number of clones	120	120
Clones analysed	51	51	Clones analysed	68	68
Positives	1	14	Positives	5	19
Negatives	50	37	Negatives	63	49
Both junctions	1		Both junctions	5	
% Gene targeting	1.96	27.45	% Gene targeting	7.35	27.9
% Gene targeting in both junctions	1.96		% Gene targeting in both junctions	7.35	

Table 10. Analysis of the efficiency of TI of the therapeutic PGK-h*FANCA* donor into the *Mbs85* locus of FA-A MEFs analysed in experiments 1 and 2 of gene targeting. Analysis of the TI in the 5’ and the 3’ integration junctions in the 54 and 120 clones generated in two independent experiments.

Fluorescence *in situ* hybridization (FISH) is a powerful technique that can be used to visualize transgene integration sites in the expected chromosome. Metaphase chromosomes were used to characterize the specific PGK-h*FANCA* donor integration in chromosome 7, where the *Mbs85* locus is located. Two sets of probes were designed. One set recognized chromosome 7, which appears as two green spots corresponding to the two chromatids of chromosome 7. Another probe recognized the PGK-h*FANCA* transgene and appears as a red spot.

As shown in metaphase 1 of FA-A MEFs (**Figure 19 A**), we observed four duplicated green spots corresponding to four chromosomes 7 (indicated with four green arrows), suggesting that FA-A MEFs were tetraploid cells. Moreover, FA-A MEFs presented a basal genomic instability that could be observed in metaphase 2 (**Figure 19 B**), showing that one of the four chromosomes 7 was longer (indicated with blue arrows) than the others, probably due to chromosome rearrangements.

In gene-edited FA-A MEF clones, we studied the targeted integration of the PGK-h*FANCA* cassette into chromosome 7. The integration of PGK-h*FANCA* donor (it appears as a red spot pointed out by a red arrow) could be observed close to the centromere of chromosome 7 (it appears as two green spots pointed out by a green arrow). The colocalization of the red spot with the green spots in the same metaphase spread indicates the targeted integration of the donor into chromosome 7, strongly suggesting targeted integration into the *Mbs85* locus. Targeted integration events observed in different samples are marked with an orange circle.

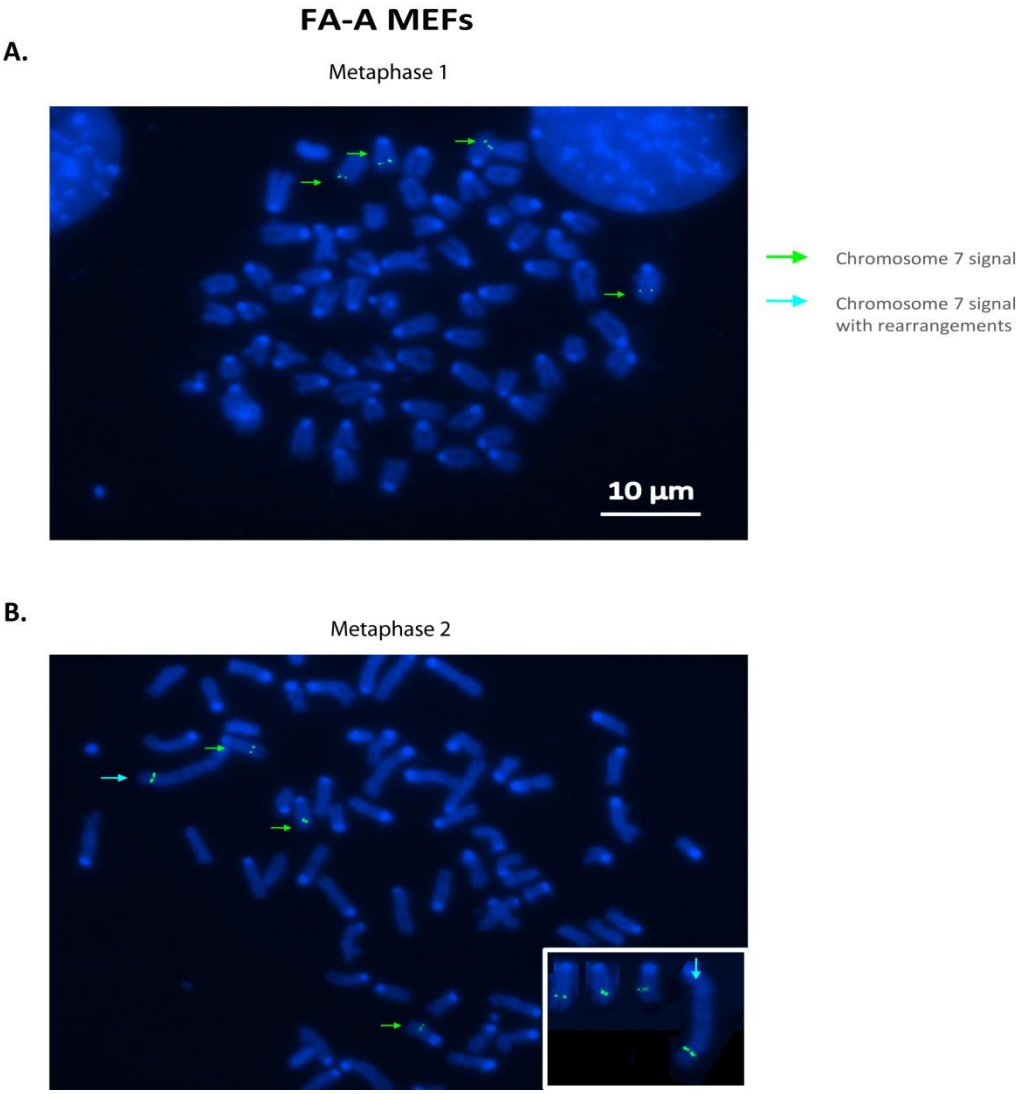


Figure 19. Representative metaphase spreads of uncorrected FA-A MEFs. Metaphase spreads used as a negative control of targeted integration in FA-A MEFs. Each chromosome 7 is indicated as a green arrow pointing out two green spots. **A)** Tetraploids FA-A MEFs (4n) without rearrangements are shown in metaphase 1. **B)** FA-A MEFs with rearrangements are observed in metaphase 2, with a long chromosome 7 indicated with a blue arrow. In a box are indicated the chromosomes 7 in a larger size. DAPI chromosome staining is shown in blue. Scale bar represents 10 µm for all the microphotographs.

In two out of the five analysed clones, #7# and #76# (**Figure 20 A and B**) we confirmed the TI of the PGK-h*FANCA* cassette in chromosome 7 (marked with an orange circle). However, random integration events were also observed in some of the metaphase spreads from clones #76#, #79# and in the bulk of gene-edited FA-A MEFs (T+D0.75). These could be recognized because red spots did not colocalize with the green spots in the same metaphase spread (**Figure 20 B and C).**

Thus, three different situations were observed in metaphase spreads from gene-edited FA-A MEF clones: metaphases in which only an on-target integration event took place (**Figure 20 A**), metaphases that presented both on-targeted and random integration events (**Figure 20 B**) and metaphases with only random integration events (**Figure 20 C**).

FISH studies confirmed the targeted integration of the therapeutic PGK-h*FANCA* donor into chromosome 7. Nevertheless, off-target integration events were also deduced in these studies.

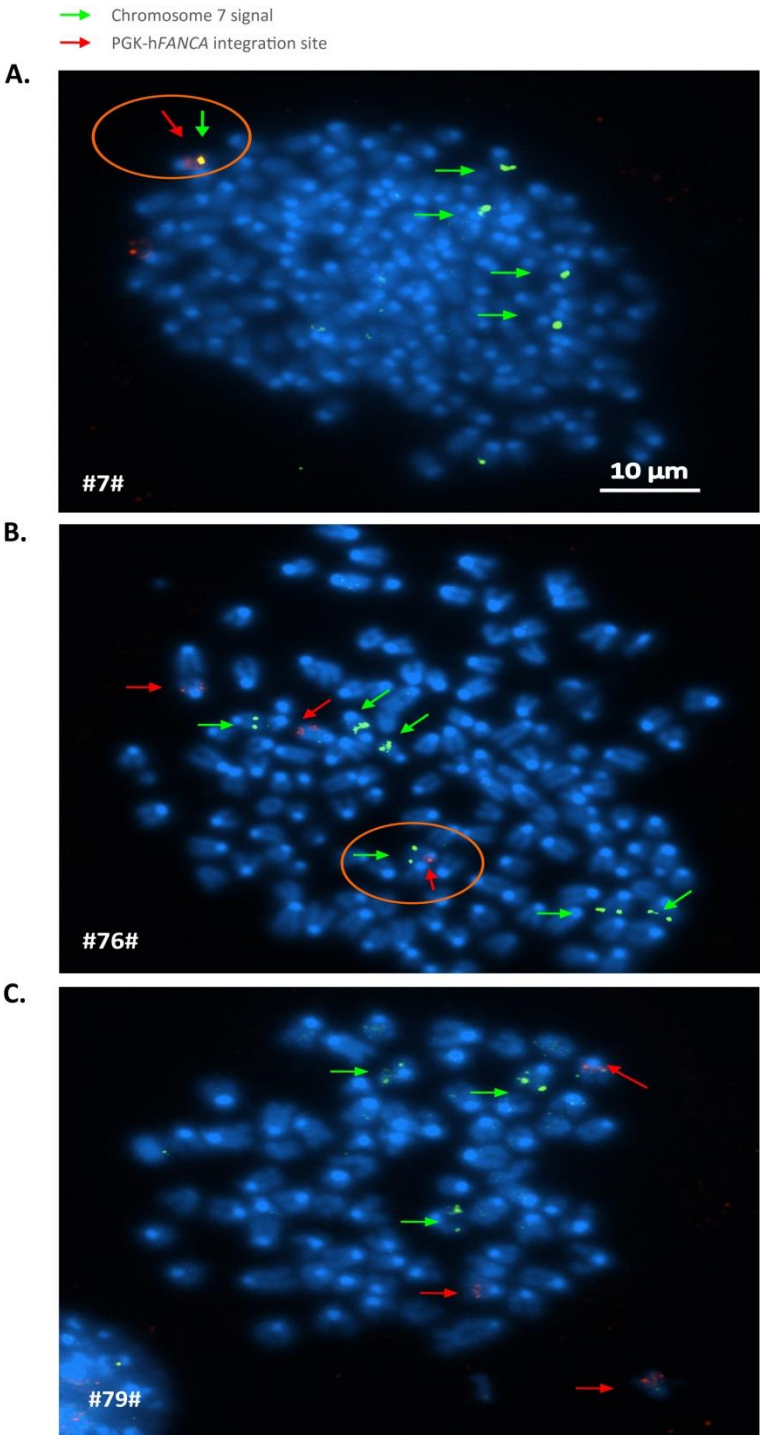


Figure 20. FISH analysis in gene-edited FA-A MEF clones to investigate the integration of the therapeutic PGK-h*FANCA* donor into chromosome 7, where the *Mbs*85 locus is located. Analysis of TI of the therapeutic PGK-h*FANCA* cassette (red spot signal and red arrows) in chromosome 7 (two green signals corresponding to the two chromatids indicated with a green arrow). DAPI chromosome staining is shown in blue. Scale bar represents 10 µm for all the microphotographs. Representative metaphase spreads of different gene-edited FA-A MEF clones: **A)** Targeted integration event in clone #7#; **B)** Targeted and random integration events in clone #76#; **C)** Random integration events in clone #79#.

2.4. Study of the phenotypic correction of gene-edited FA-A MEF clones

Once the integration of the therapeutic PGK-h*FANCA* donor into the *Mbs85* locus was confirmed in different FA-A MEF clones, the functionality of the cassette was first determined by the analysis of hFANCA expression. Analyses were made by WB both in gene-edited clones and in the bulk population of gene-edited FA-A MEFs, either unselected or after MMC selection. Since mouse MEFs do not express the human FANCA protein, we normalized the expression of hFANCA using a healthy human donor lymphoblastic cell line (HD LCL). As shown in **Figure 21**, neither FA-A or WT MEFs nor FA-A LCLs expressed detectable levels of the hFANCA protein. However, we observed hFANCA expression both in the bulk of gene-edited FA-A MEFs (T+D0.75), and in gene-edited FA-A MEF clones (indicated in purple: #7# and #77# in **Figure 21 A**; #76# and #79# in **Figure 21 B**; and #83# in **Figure 21 C**). This result demonstrates the efficiency of our gene-editing approach to promote the expression of the therapeutic h*FANCA* gene in FA-A MEFs.

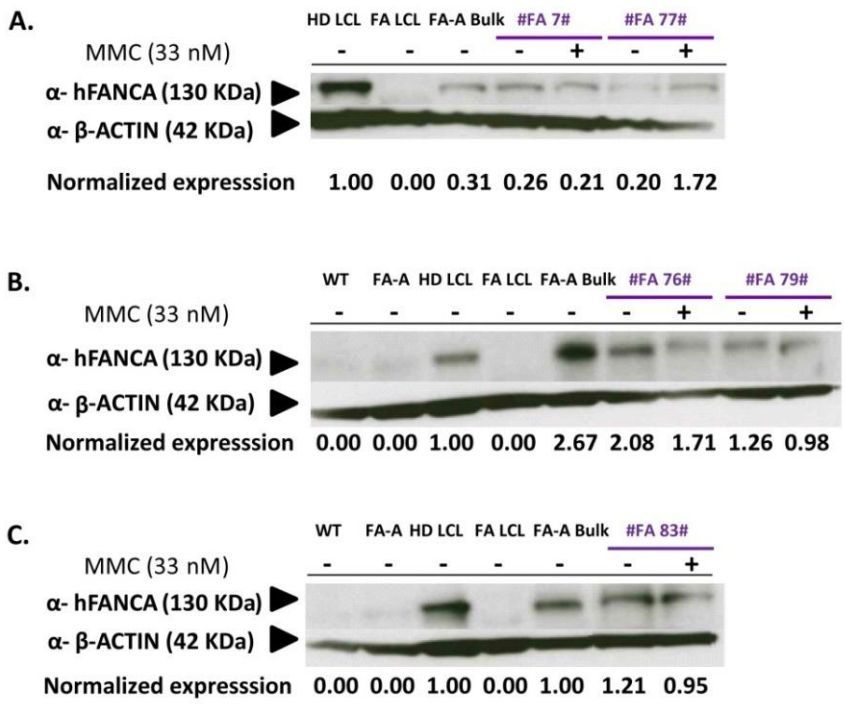


Figure 21. Western blot analysis of the hFANCA expression in gene-edited FA-A MEF clones. Analysis of hFANCA expression in the bulk of gene-edited FA-A MEFs, as well as in the gene-edited derived clones. Expression levels were calculated as a fold change with respect to β-ACTIN that was used as a loading control and then normalized against the hFANCA expression of lymphoblastic cells derived from a healthy human donor (HD LCL). FA LCL: lymphoblastic cells derived from a Fanconi anemia A patient; FA-A: immortalized non-gene-edited FA-A MEFs; WT: immortalized non-gene-edited WT MEFs; FA-A Bulk: bulk of gene-edited FA-A MEFs (T+D0.75). The clones analysed are coloured in purple: #7#, #77# **(A)**, #76#, #79# **(B)** and #83# **(C)**, unselected or after 33 nM of MMC selection.

Second, as FA cells are characterized by their hypersensitivity to DNA interstrand cross-linking agents such as mitomycin C (MMC), the MMC sensitivity of uncorrected and gene-edited FA-A MEFs was tested. Cells were cultured with different doses of MMC and their survival was analysed at 14 days post-treatment. Gene-edited clones #7#, #76# and #77# showed increased MMC resistance as compared to non-gene-edited FA-A MEFs. Interestingly,

in clones #7# and #76# the presence of a single integration in chromosome 7 was confirmed by FISH. Strikingly, clones #79# and #83# did not show correction of their MMC hypersensitivity, even though, a significant expression of hFANCA was observed in these clones (**Figure 22**).

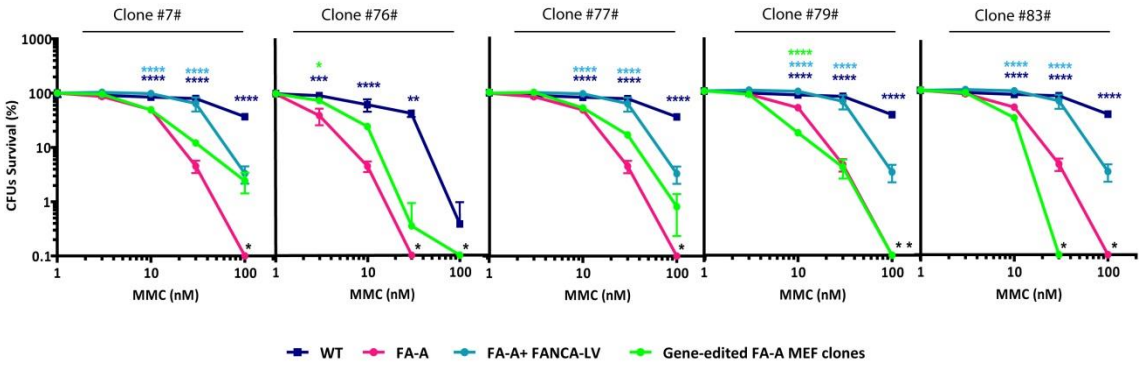


Figure 22. MMC survival curve of gene-edited FA-A MEFs. Survival curves after exposure to different concentrations of MMC (0, 3, 10, 30, 100 nM) of gene-edited FA-A MEF clones. Representative MMC sensitivity curves of five gene-edited clones (#7#; #76#; #77#; #79# and #83#). In each curve the analysed clone is represented, together with untransfected WT and FA-A MEFs, and FA-A MEFs transduced with a PGK-*hFANCA*-wPRE* LV at MOI 1. Black asterisk (*) indicates that no colonies were generated in this condition at the corresponding doses of MMC. The survival shown would correspond to the growth of a single colony in these cultures. (*) P-value <0.05, (**) P-value <0.01, (***) P-value <0.001, (****) P-value <0.0001 indicate significant differences with respect to FA-A group, with a two-way ANOVA followed by a post-hoc Bonferroni test. The coloured asterisks indicate groups with differences with respect to the FA-A group.

Finally, we studied the chromosomal instability induced by MMC, both in WT MEFs, in uncorrected FA-A MEFs, and in gene-edited FA-A MEF clones. For that purpose, cells were treated for 24 hours with 40 nM MMC and their metaphases were studied. First, we observed that gene-edited FA-A MEF clones presented a similar percentage of aberrant cells as compared with non-gene-edited FA-A MEFs (**Figure 23 A**), most probably because of the characteristic genetic instability of these cells. However, the number of MMC-induced aberrations per chromosome in FA-A gene-edited clones was 11-40 fold lower in comparison with non-gene-edited FA-A MEFs (**Figure 23 B**).

These results confirm the functionality of the therapeutic cassette both by the expression of the human FANCA protein, and by the reversion of the characteristic hypersensitivity of FA-A cells to MMC.

Altogether these studies demonstrate that the specific integration of the PGK-h*FANCA* transgene into the *Mbs85* locus confers a phenotypic correction to FA-A MEFs.

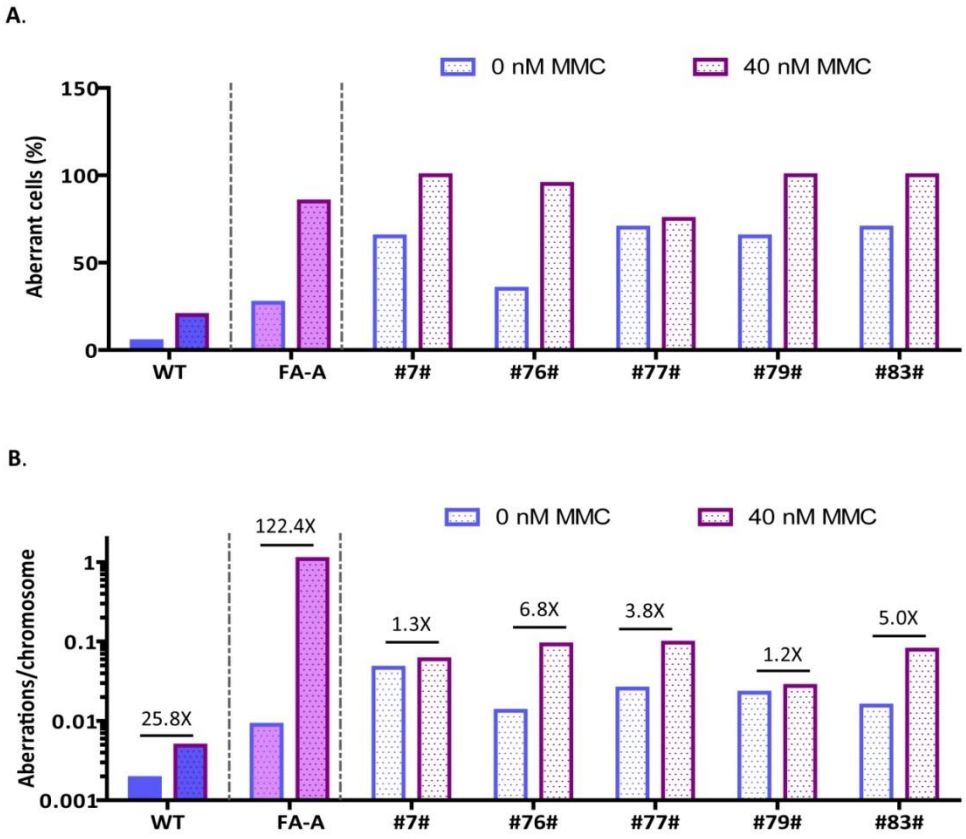


Figure 23. Chromosomal instability study in gene-edited FA-A MEF clones. Chromosomal aberrations analysed in metaphases of gene-edited FA-A MEF clones #7#, #76#, #77#, #79# and #83, in comparison with non-gene-edited WT and FA-A MEFs. **A)** Percentage of aberrant cells treated or not with mitomycin C. **B)** Chromosomal aberrations per chromosome in samples treated or not with mitomycin C. The fold change between samples treated with mitomycin C and non-treated ones is also indicated.

3. GENE-TARGETING STUDIES IN PRIMARY MOUSE HEMATOPOIETIC PROGENITORS

3.1. Optimization of the nucleofection of plasmid DNA and mRNA required for gene targeting of mouse hematopoietic progenitors

Once the targeted integration of the therapeutic PGK-h*FANCA* donor into the *Mbs85* locus of FA-A MEFs was demonstrated, we aimed at investigating the feasibility of conducting a similar approach in mouse hematopoietic stem and progenitor cells (mHSPCs). Because there are no published data about procedures capable of efficiently nucleofecting mHSPCs, we first aimed at optimizing the nucleofection of mouse Lin^- BM cells with plasmid DNA.

Lin^- BM cells were purified by cell sorting and then prestimulated for 48 hours before nucleofection to facilitate mHSPCs to enter in cell-cycle, a requirement to promote HDR. Between $1x10^6$ to $2.3x10^6$ cells were nucleofected either with 2 µg of an *EGFP* control plasmid or of a pMAX *EGFP*® control plasmid. The viability (percentage of $DAPI^-$ cells) of Lin^- cells and the transfection efficiency in these cells (percentage of $EGFP^+$ $DAPI^-$ cells) were evaluated at 48 hours post-nucleofection.

Several programs were tested with two Lonza nucleofector devices (I and 4D). We obtained better results with the 4D-nucleofector in comparison with the nucleofector device I (experiments Lin^- 1 and 2 were done with nucleofector I, **Table S4**). Then, different nucleofection programs of the 4D-nucleofector were tested: EO-100, ED-113, DT-113, (experiments Lin^- 3 and 4) and DN-110, EA-105, ED-123, DS-123 and EN-113 (experiments Lin^- 25 and 26). In these four experiments, the viabilities obtained at 48 hours post-nucleofection ranged from 0.8% to 9%, and the transfection efficiencies from 3% to 71% (**Figure 24 A**).

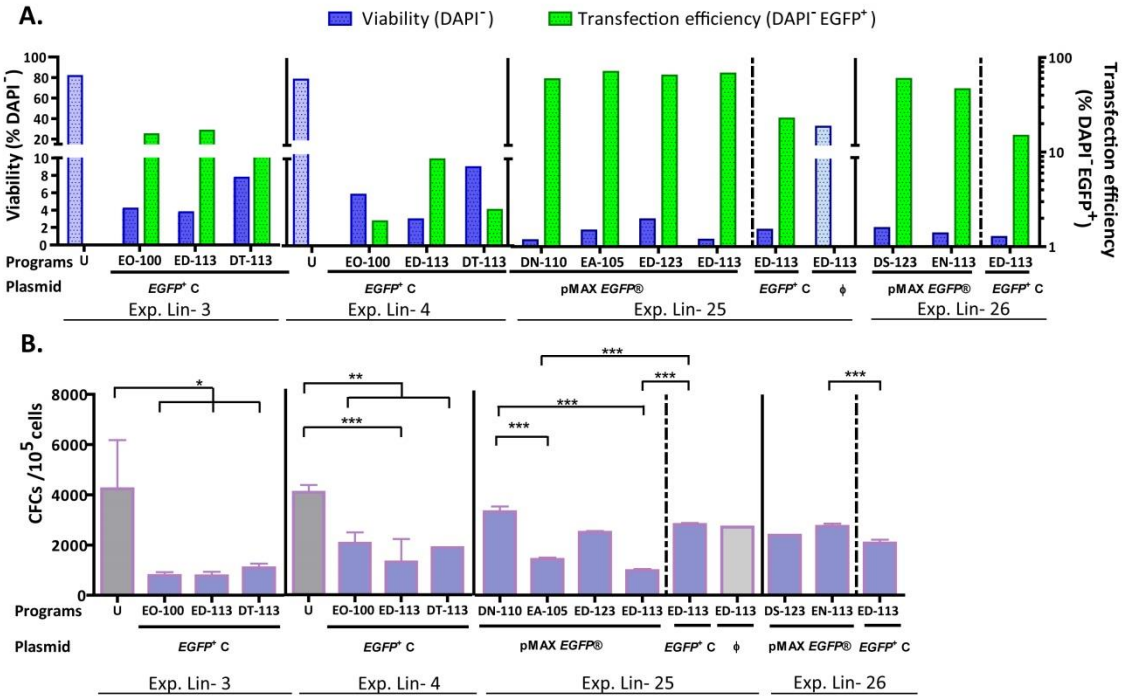


Figure 24. Setting up the nucleofection conditions in Lin⁻ BM cells. A) Evaluation of viability (percentage of total DAPI⁻ cells) and transfection efficiency (percentage of EGFP⁺DAPI⁻ cells) at 48 hours post-nucleofection. **B)** Clonogenic assays to evaluate the ability of nucleofected Lin⁻ BM cells to generate hematopoietic colonies. Four different experiments (Lin⁻ 3-4, and Lin⁻ 25-26, **Table S4**) were performed with different programs: EO-100, ED-113, DT-113, DN-110, EA-105, ED-123, DS-123, and EN-113. All the cells were nucleofected with the *EGFP* control plasmid or with the pMAX *EGFP*® control provided by Lonza; U: untransfected cells; ϕ: transfected cells without DNA (mock condition). The different experiments are separated with black lines. Dashed lines separate the different *EGFP* control plasmids used. Data are the mean ± S.D. (n=3 experimental replicates of one experiment). Statistical differences were calculated for each of the experiments. (*) P-value <0.05, (**) P-value <0.01, (***) P-value <0.001 indicate significant differences with a one-way ANOVA followed by a *post-hoc* Tukey test, or non-parametric Kruskal-Wallis and median test followed by a *post-hoc* Dunn's test.

To investigate the clonogenic potential of nucleofected Lin⁻ BM cells, cells were seeded in methylcellulose and colonies were scored seven days later. DNA nucleofection normally resulted in a 2.8 to 5.4-fold decrease in the number of colonies with respect to untransfected or transfected cells without DNA (mock cells). With the exception of the EA-105 and the ED-113 programs nucleofected with the pMAX *EGFP*® control plasmid, most of programs generated approximately the same number of colonies, with no marked differences among them (**Figure 24 B**). Taking into account the efficiencies of transfection, viability and clonogenic ability of nucleofected Lin⁻ BM cells observed in the different experiments conducted, we decided that the ED-113 program of nucleofection was the most efficient in nucleofecting mouse HSPCs.

In another experiment we also tested the response to nucleofection of Lin⁻ cells that express high levels of markers such as Ly6 E/A/ Sca-1 and C-Kit, and that are known as LSK cells (Lin⁻ Sca-1⁺ c-Kit⁺). These cells are enriched in mHSPCs with multi-lineage and long-term reconstitution activity. In this experiment, 0.4 µg of the *EGFP* control plasmid were used to nucleofect $1x10^6$ cells using the ED-113 program. We obtained 4% transfection efficiency in the total population of nucleofected Lin⁻ cells (percentage of EGFP⁺ DAPI⁻ cells) with 13% viability

(percentage of DAPI$^-$ cells). Moreover, the percentage of LSK$^+$ cells in the live-population was 3.5% (percentage of LSK$^+$DAPI$^-$ cells), this percentage was 4.4-fold lower than the percentage of LSK$^+$ cells observed in the mock samples. Thus, the nucleofection of plasmid DNA not only affected the viability of the hematopoietic colonies but also the percentage of LSK hematopoietic progenitors. Additionally, the percentage of LSKs expressing EGFP was 3.8% (percentage of EGFP$^+$LSK$^+$DAPI$^-$ cells) and the LSK$^+$ cells that were positive for EGFP in the whole population were only 0.14% (percentage of EGFP$^+$LSK$^+$ cells).

These experiments showed that DNA nucleofection was notable toxic in mouse HSPCs. mRNA transfection is directly delivered to the cytoplasm and available for translation, and it has been proposed to be more efficient than DNA for the delivery of nucleases. Thus, we tested the nucleofection of mRNA in Lin$^-$ BM cells. Four experiments were performed in which ED-113, DS-118, and DK-100 programs were tested to evaluate the viability and the transfection efficiency of the procedure.

First, two different experiments (Lin$^-$ 16 and Lin$^-$ 24, see **Table S4** for more details of the experiments) were performed with the ED-113 program. In these experiments, two different *EGFP* control mRNA batches were tested (referred to as mRNA *EGFP* (A) and mRNA *EGFP* (B), and reviewed in Materials & Methods, section 1.3.2). The efficiency of all mRNA batches used in these experiments had been previously confirmed in nucleofection experiments with CD34$^+$ cells. In these experiments, 0.8-1.4x10^6 Lin$^-$ BM cells were nucleofected with eight different doses (from 2.5 to 50 µg) of any of these mRNAs. The viability and transfection efficiency were analysed at different time points post-transfection (24, 48, 72 and 96 hours). At any of the time points analysed, with mRNA the viabilities were higher than as with 2.5 µg of plasmid DNA (in the range of 8% to 64% with mRNA as compared to 6% with DNA), but the transfection efficiencies were always much lower (in the range of 0-0.2% with mRNA as compared to 28% with DNA).

Furthermore, two additional experiments (experiments Lin$^-$ 6 and Lin$^-$ 40, **Table S4**) were performed with DS-118 and DK-100 programs (**Figure 25**). With the DS-118 program, 0.5x10^6 cells were nucleofected with two doses (0.4 and 2 µg) of an *EGFP* control mRNA (A) and with 0.4 µg of a control DNA. With the DK-100 program, 1x10^6 cells were nucleofected with 4 µg of mRNA and 2.5 µg of a control DNA. In this last experiment two different mRNA were tested: the commercial *StemMACS Nuclear EGFP mRNA,* (mRNA *EGFP* (C)), and another batch of mRNA *EGFP* (B). Viability and transfection efficiency were analysed at 24 and 48 hours post-nucleofection. mRNA nucleofection maintained cell viability (it ranged from 11.9% to 40.1%). However, the transfection efficiency was again very low (below 1.6%) as compared to DNA (from 3.3% to 33.3%) (**Figure 25 A**). With respect to the ability of the Lin$^-$ BM cells to generate hematopoietic colonies, their clonogenic ability was from 2.7 to 67-fold higher when the cells were nucleofected with mRNA in comparison with those nucleofected with DNA (**Figure 25 B**).

Taking into account all the studies performed in Lin$^-$ mHSPCs with plasmid DNA and mRNA, we conclude that DNA nucleofection with the ED-113 program constitutes the most efficient procedure of the tested ones to transfect mouse HSPCs in spite of its significant toxicity.

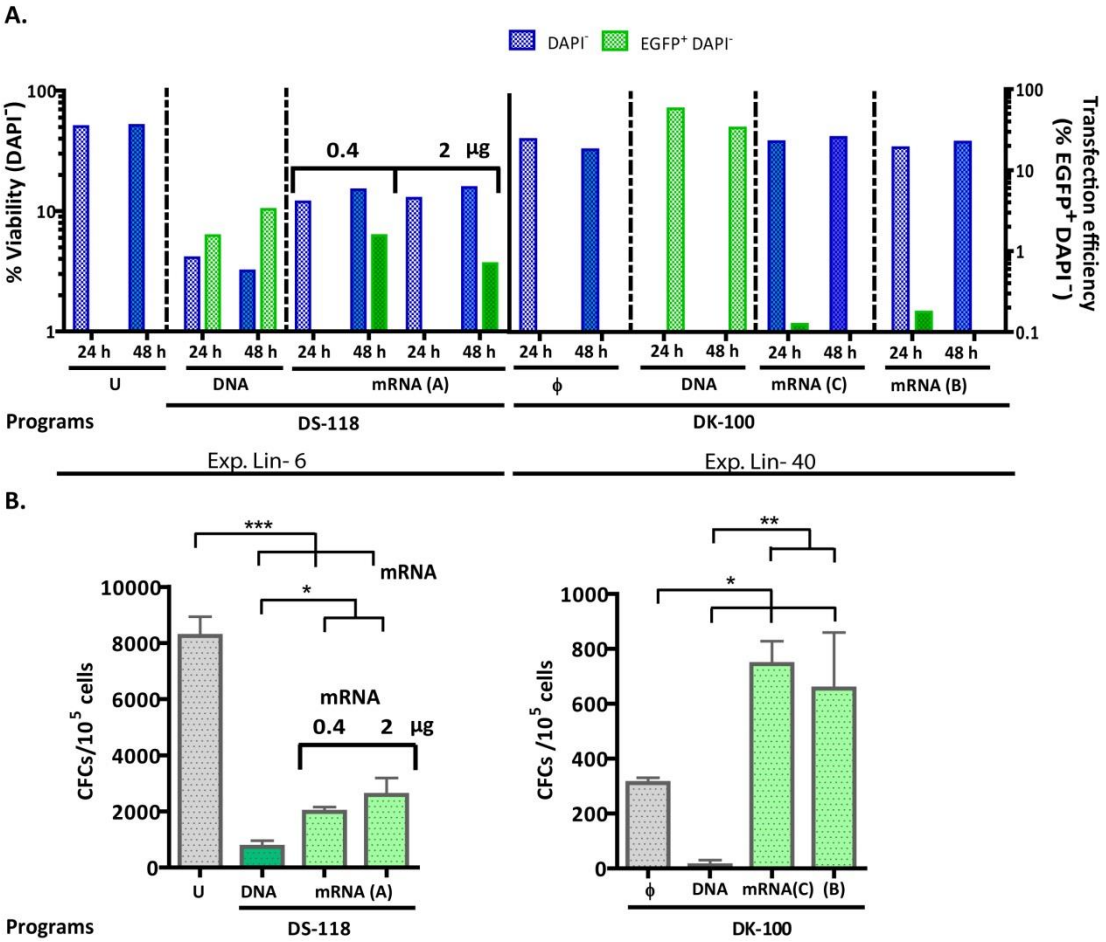


Figure 25. mRNA nucleofection in Lin⁻ BM cells. **A)** Evaluation of the viability (percentage of total DAPI⁻ cells) and transfection efficiency (percentage of EGFP⁺ DAPI⁻ cells) with DS-118 and DK-100 nucleofection programs. With the DS-118 program, two doses of the mRNA (A) were tested: 0.4 and 2 µg. With the DK-100 program, one dose of 4 µg was tested with two batches of mRNA, the *StemMACS Nuclear EGFP* mRNA (C) and mRNA *EGFP* (B). Data referred to are from experiments Lin⁻ 6 and Lin⁻ 40 (**Table S4**). The different experiments are separated with black lines. Dashed lines separate the different conditions. **B)** Clonogenic assays to evaluate the ability of the Lin⁻ BM cells to generate hematopoietic colonies under the conditions shown in A). Data are the mean ± S.D. (n=3 experimental replicates from one experiment). (*) P-value <0.05, (**) P-value <0.01, (***) P-value <0.001 indicate significant differences with a non-parametric Kruskal-Wallis and median test followed by a *post-hoc* Dunn's test.

3.2. Evaluation of mTALEN cleavage efficacy in the *Mbs85* locus of WT mouse HSPCs

Once nucleofection conditions were established in Lin⁻ mouse BM cells with plasmid DNA, we evaluated by Surveyor assay the cleavage efficacy of the NN-mTALEN pair in the *Mbs85* locus of mHSPCs from WT mice (FVB strain) as shown in **Figure 26**. The cleavage efficacies of the NN-mTALEN pair observed in three independent experiments were 30.41 ± 5.95 % in WT Lin⁻ BM cells.

These results are consistent with those obtained in FA-A MEFs and demonstrate the efficient cleavage of the NN-mTALEN pair in the *Mbs85* locus of WT Lin⁻ BM cells.

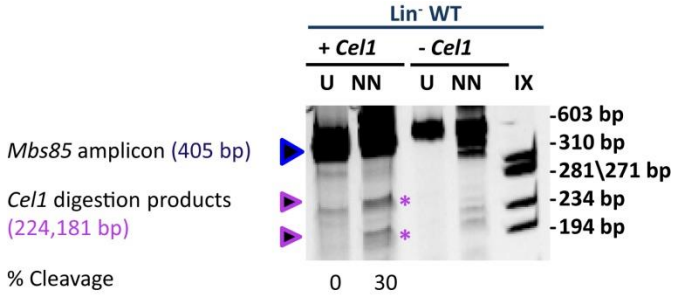


Figure 26. Evaluation of the cleavage efficacy of mTALEN in the *Mbs85* locus of WT Lin⁻ BM cells. Representative electrophoresis gel representing the disruption of the target locus using 2.5 µg of each NN-mTALEN monomer indicated as NN. Samples not digested with the endonuclease were used as controls. The extent of cleavage measured as the mean percentage of modified alleles is indicated. Arrows indicate the size of the parental band (405 bp, in blue) and the expected positions of the digestion products (224 bp and 181 bp, in purple), that are also indicated with purple asterisks. U: untransfected cells; IX: DNA molecular weight marker.

3.3. Gene-targeting analyses in WT hematopoietic progenitors with the PGK-*EGFP* reporter donor

Once we had demonstrated the NN-mTALEN efficacy in cleaving the *Mbs85* locus of WT Lin⁻ BM cells, gene-targeting experiments were first performed in these cells with the aim of investigating whether the same approach previously conducted in FA-A MEFs was feasible in WT mHSPCs. For that purpose, a reporter donor template, the PGK-*EGFP*-SV40pA donor was constructed.

3.3.1. Evaluation of viability, EGFP expression and clonogenic ability of nucleofected WT Lin⁻ BM cells

In these experiments, two doses of the NN-mTALEN pair, 0.75 µg and 2.5 µg of each monomer were nucleofected together with 2 or 4 µg of the PGK-*EGFP* reporter donor, both constructs as plasmid DNA.

With respect to the viability of nucleofected Lin⁻ BM cells measured at 48 hours post-nucleofection, DNA-transfected cells showed a 7-fold decrease in the viability either with respect to untransfected or to mock transfected cells (**Figure 27 A**). Among the conditions transfected with DNA no significant differences were observed. The transfection efficiency of these experiments was analysed using 2.5 µg of an *EGFP* control plasmid and resulted in 22.65 ± 14.07% of $EGFP^+$ cells. When cells were nucleofected with different doses of the mTALEN and the PGK-*EGFP* reporter donor, the percentage of $EGFP^+$ cells at 48 hours post-nucleofection ranged from 11.7% to 27.6%, regardless of the doses of mTALEN. At this time, cells that were nucleofected only with the donor also expressed EGFP in a range that varied between 3% and 22% (**Figure 27 A**). In both conditions, the expression of EGFP was expected due to the expression of the episomal donor construct plasmid DNA. Moreover, no significant differences were observed regarding the mean fluorescence intensity (MFI) comparing the samples nucleofected with only the donor with those nucleofected with the mTALEN and the donor (**Figure 27 B**).

The ability of the WT nucleofected Lin⁻ BM cells to generate hematopoietic colonies was also evaluated in these experiments. DNA nucleofection reduced the ability of BM

hematopoietic progenitors to generate colonies approximately 9-fold with respect to mock transfected cells (**Figure 27 C**). However, we did not find significant differences among conditions nucleofected with the donor, regardless of the presence of the TALEN, indicating that the different doses of TALEN did not add further toxicity to the process.

These results show a significant EGFP expression in gene-edited WT Lin$^-$ BM cells at 48 hours post-nucleofection either with the mTALEN and the reporter donor or with only this donor. Moreover, the viability and the clonogenic ability of nucleofected Lin$^-$ BM cells were markedly decreased by plasmid DNA nucleofection.

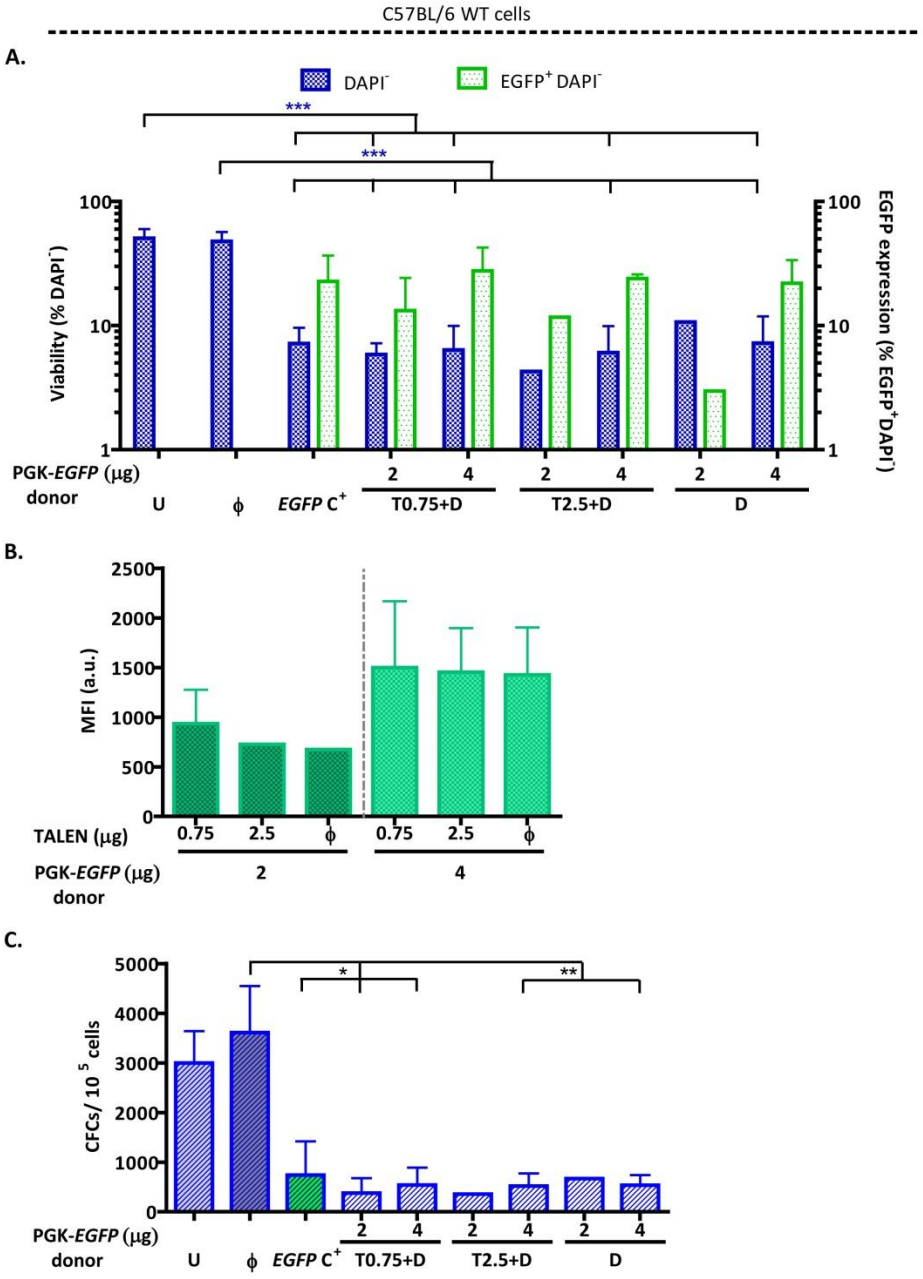


Figure 27. Evaluation of viability, efficiency of EGFP expression and clonogenic ability of WT Lin$^-$ BM cells nucleofected with the mTALEN and the PGK-*EGFP* reporter donor at 48 hours post-nucleofection. A) Analysis of viability (percentage of DAPI$^-$ cells) and percentage of EGFP$^+$ cells determined at 48 hours post-nucleofection. U: untransfected cells; ϕ: transfected cells without DNA (mock condition); *EGFP* C$^+$:

EGFP control plasmid; T: different doses of NN-mTALE nuclease monomers (from 0.75 to 2.5 µg); T+D: 0.75 or 2.5 µg of each NN-mTALE nuclease monomer plus the indicated doses of the PGK-*EGFP* reporter donor (2 or 4 µg); D: 2 or 4 µg of the PGK-*EGFP* reporter donor. Data are the mean ± S.D. (n=5 experiments). (***) P-value <0.001 indicates significant differences with a one-way ANOVA followed by a *post-hoc* Tukey test. **B)** Geometric Mean Fluorescent Intensity (measured in arbitrary units, a.u.) of the indicated conditions at 48 hours post-nucleofection. ϕ: transfected cells with the donor without mTALEN. Data are the mean ± S.D. (n=5 experiments). No statistical differences were found among groups with the parametric one-way ANOVA test. **C)** Clonogenic assay to evaluate the ability of the Lin$^-$ BM cells to generate hematopoietic colonies under the conditions shown in A). Data are the mean ± S.D. (n=5 experiments). (*) P-value <0.05, (**) P-value <0.01 indicate significant differences with a one-way ANOVA followed by a *post-hoc* Tukey test. Data referred to are from experiments Lin$^-$ 27-34 (**Table S4**).

Previous experiments of our laboratory indicated that the viability of nucleofected cells was recovered when cells were maintained between 4 to 6 days in culture after nucleofection. Thus, a new set of experiments was conducted maintaining the nucleofected Lin$^-$ BM cells for 5 days instead of 2 days in culture before performing different analyses.

At 5 days post-nucleofection, the viability in nucleofected samples with plasmid DNA was in the same range with respect to data obtained at 2 days post-nucleofection (ranging from 4% to 13%) (**Figure 28 A** and **Figure 27 A**). The EGFP expression in the DNA nucleofected samples was lower at 5 days post-nucleofection (it ranged from 0.69% to 0.81%) in comparison with 2 days post-transfection (it ranged from 2.97% to 27.57%), something that was expected due to the dilution of the episomal plasmid in these cells (**Figure 28 A** and **Figure 27 A**). In clonogenic cultures, we observed a trend indicating that cells maintained in culture for 5 days had higher clonogenic ability as compared to those maintained in culture only for 2 days (**Figure 28 B**). However, since no significant differences were observed in any of the analysed parameters, the time point of 2 days post-nucleofection was chosen for further analyses.

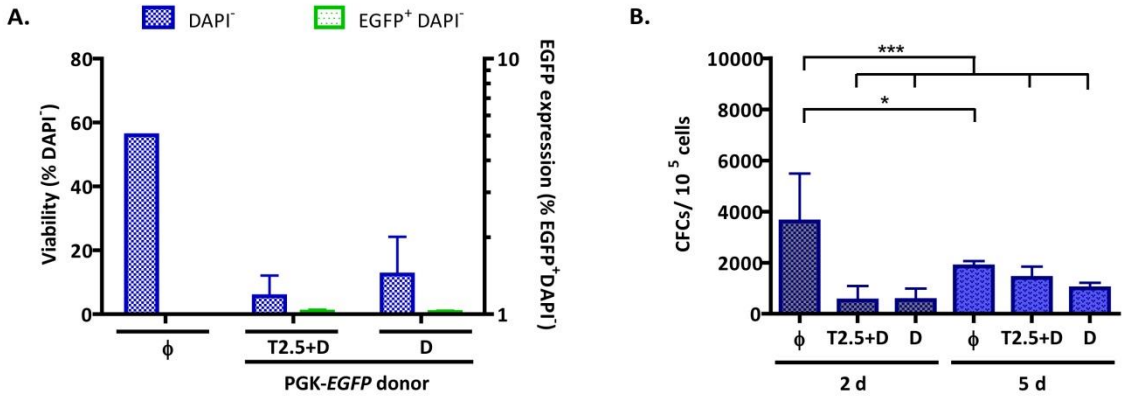


Figure 28. Evaluation of the viability, efficiency of EGFP expression and clonogenic ability of nucleofected WT Lin$^-$ BM cells maintained for 2 and 5 days post-nucleofection. ϕ: transfected cells without DNA (mock condition); T2.5+D: 2.5 µg of each NN-mTALE nuclease monomer plus 4 µg of the PGK-*EGFP* reporter donor; D: 4 µg of the PGK-*EGFP* reporter donor. **A)** Analysis of viability (percentage of DAPI$^-$ cells) and percentage of EGFP$^+$ cells in the different conditions at day 5 post-nucleofection. Data are the mean ± S.D. (n=3 experiments). No statistical differences were found among groups with a non-parametric Kruskal-Wallis and median test followed by a *post-hoc* Dunn's test. **B)** Clonogenic assay to evaluate the ability of the nucleofected Lin$^-$ BM cells to generate hematopoietic colonies under the conditions shown in A). Data are the mean ± S.D. n=5 experiments (2 days expansion) and n=3 experiments (5 days in expansion). (*) P-value <0.05, (***) P-value <0.001 indicate significant differences with a one-way ANOVA followed by a *post-hoc* Tukey test. Data referred to are from experiments Lin$^-$ 35 & Lin$^-$ 37-38 (5 days post-nucleofection) (**Table S4**).

In a next set of experiments, we investigated the viability and EGFP$^+$ expression in cells maintained in culture for 14 days post-nucleofection. When both the mTALEN and the donor were nucleofected, the percentage of cells that remained positive for EGFP expression ranged from 0.12% (with 0.75 µg of mTALEN and 4 µg of donor) to 0.28% (with 2.5 µg of mTALEN and 4 µg of donor). However, when only the donor was nucleofected (4 µg), no EGFP$^+$ cells were detected at 14 days post-nucleofection (**Figure 29 A and B**). These results suggest the generation of a low, though significant, number of gene-edited mHSPCs, when both the mTALEN and the donor were nucleofected in Lin$^-$ BM cells.

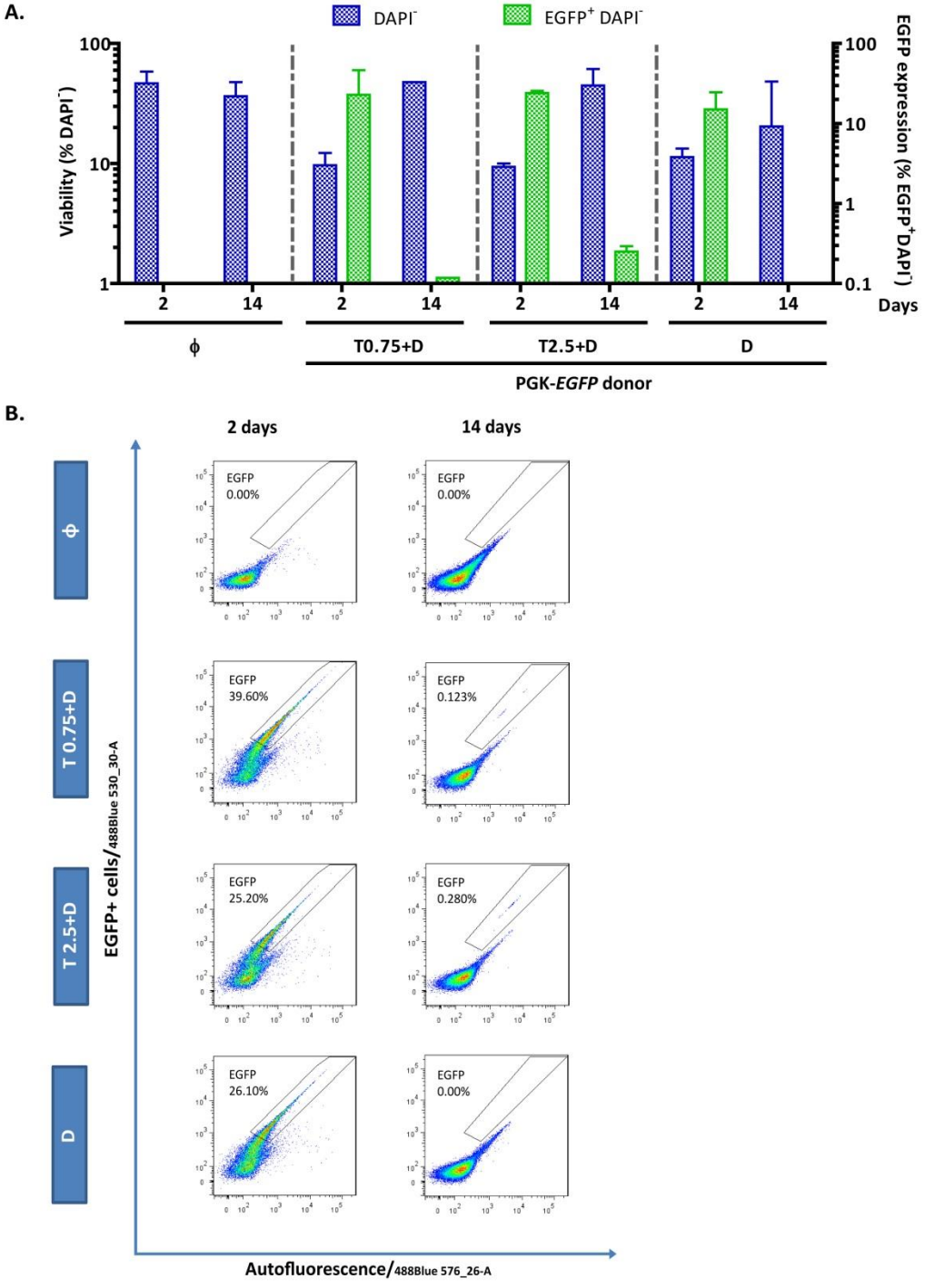


Figure 29. Evaluation of the viability and efficiency of EGFP expression of WT Lin$^-$ BM cells nucleofected with the mTALEN and the PGK-*EGFP* reporter donor at days 2 and 14 post-nucleofection.

A) Analysis of viability (percentage of DAPI$^-$ cells) and percentage of EGFP$^+$ cells in the different conditions at 2 and 14 days post-nucleofection. ϕ: transfected cells without DNA (mock condition); T+D: 0.75 or 2.5 µg of each NN-mTALE nuclease monomer plus 4 µg of the PGK-*EGFP* reporter donor; D: 4 µg of the PGK-*EGFP* reporter donor. Data are the mean ± S.D. (n=2-3 experiments). Statistical analysis could not be performed. **B)** Representative flow cytometry dot plots of EGFP$^+$ cells analysed at day 2 and 14 post-nucleofection in nucleofected Lin$^-$ BM cells from conditions mentioned in A). EGFP$^+$ expression was determined discarding autofluorescent cells (described in Materials & Methods section 4.2.1).

Figure 30 A shows the presence of EGFP$^+$ cells maintained in liquid culture at 14 days post-nucleofection. The MFI was also analysed at 14 days post-nucleofection, and compared with data obtained at 2 days post-nucleofection. As shown in **Figure 30 B**, significantly higher MFIs were observed at 14 days compared with 2 days post-nucleofection in samples nucleofected both with the mTALEN and the donor. These results suggest that the integration of the PGK-*EGFP* cassette increased the expression of the EGFP transgene (**Figure 30 B**).

Thus, the higher expression of EGFP in gene-edited samples transfected with the mTALEN designed to target the *Mbs85* locus, plus the PGK-*EGFP* donor suggest that integration events result in a higher expression of the transgene at 14 days post-nucleofection as compared to non-integrated copies at 2 days post-nucleofection.

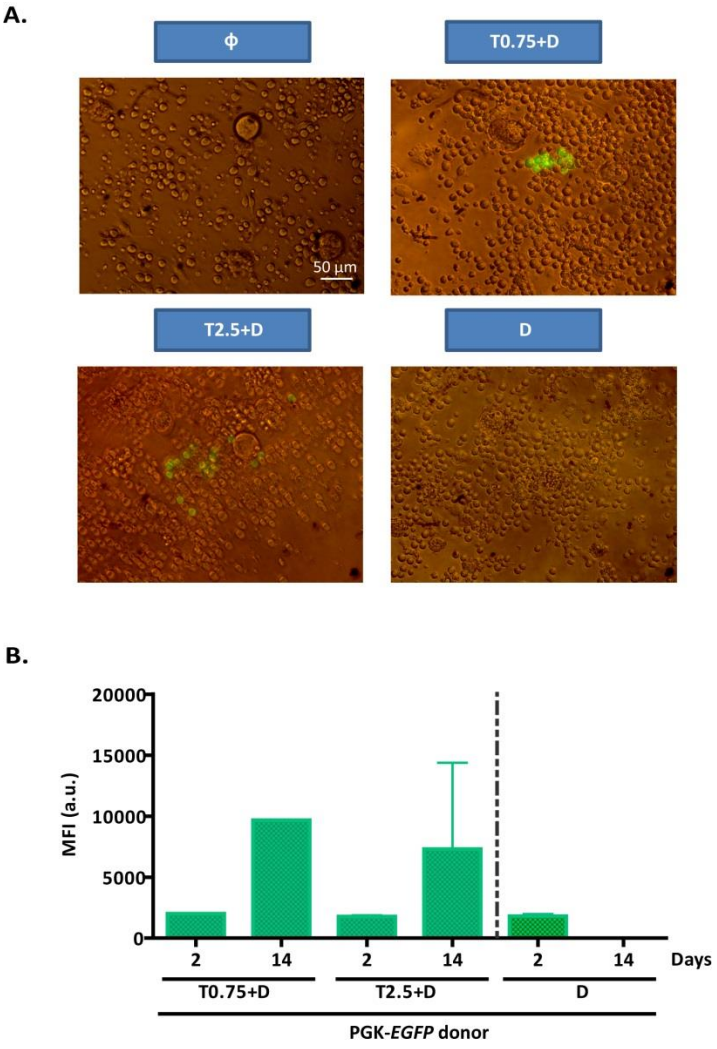


Figure 30. Analysis of EGFP expression in nucleofected WT Lin$^-$ BM cells at 14 days post-nucleofection. **A)** Representative image in liquid culture of EGFP$^+$ Lin$^-$ BM cells at 14 days post-nucleofection. EGFP$^+$

fluorescent cells could be observed in the microscope in T+D conditions. ϕ: transfected cells without DNA (mock condition); T+D: 0.75 or 2.5 µg of each NN-mTALE nuclease monomer plus 4 µg of the PGK-*EGFP* reporter donor; D: 4 µg of the PGK-*EGFP* reporter donor. Scale bar represents 50 µm for all the microphotographs. **B)** Geometric Mean Fluorescent Intensity (measured in arbitrary units, a.u.) of nucleofected cells analysed at days 2 and 14 post-nucleofection. T+D: 0.75 or 2.5 µg of each NN-mTALE nuclease monomer plus 4 µg of the PGK-*EGFP* reporter donor; D: 4 µg of the PGK-*EGFP* reporter donor. Data are the mean ± S.D. (n=2-3 experiments). Statistical analysis could not be performed.

3.3.2. Analysis of the targeted integration efficiency of the PGK-*EGFP* reporter donor into the *Mbs85* locus of WT Lin$^-$ BM cells by PCR

In these experiments, clonogenic assays of nucleofected Lin$^-$ BM cells were established. The number of hematopoietic colonies, as well as the presence of green fluorescent colonies was scored. Strikingly, no EGFP$^+$ colonies were detected. Therefore, most of the colonies were picked and analysed by PCR to evaluate the presence of the donor integration into the *Mbs85* locus.

Targeted integration frequencies obtained from eight experiments in which different doses of mTALEN (0.75 and 2.5 µg) and donor (2 and 4 µg) were used are shown in **Figure 31** (Data referred to are from experiments Lin$^-$27-Lin$^-$ 39, **Table S4**). TI frequencies were calculated from PCR analyses that were positive either for the 5' integration junction or the 3' integration junction. When 0.75 µg of each mTALEN monomer and 2 µg of donor were used, TI frequencies corresponding to the mean ± S.D. were 8.13 ± 2.65% and 2.5 ± 3.54% for the 5' and 3' integration junction analyses, respectively. When 4 µg of donor were used together with 0.75 µg of each mTALEN monomer, TI frequencies were 11.39 ± 9.87% and 6.47 ± 6.02% for the 5' and the 3' integration junctions, respectively. When the mTALEN dose was increased to 2.5 µg of each mTALEN monomer, and 2 µg of donor were used, TI frequencies resulted in 15.63% for both integration junction analyses. Finally, when 4 µg of donor together with 2.5 µg of each mTALEN monomer were used, TI frequencies of 9.83 ± 5.03% and 20.98 ± 17.60% were achieved in the 5' and the 3' integration junction analyses. Although no significant differences were found among the different conditions, the highest integration frequencies were obtained using 2.5 µg of each mTALEN monomer and 4 µg of donor. In most colonies the integration was only detected in the 5' or in the 3' integration junctions. However, in 5 out of the 8 experiments, we observed colonies that were positive for both junctions. The TI values deduced from colonies that were positive to both integration junctions ranged from 2.00 to 10.94% depending on the experiment. Additionally, in none of these experiments did we observe specific amplification of either the 5' or the 3' integration junctions in samples transfected with only the donor.

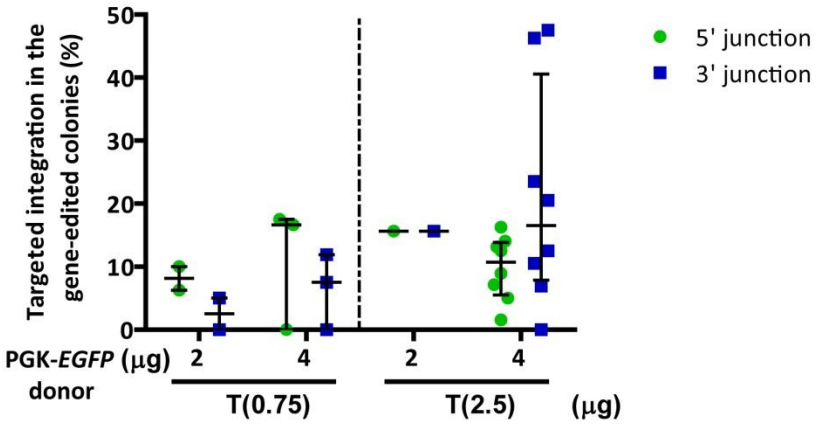


Figure 31. Targeted integration percentage of the PGK-*EGFP* reporter donor into the *Mbs85* locus of WT Lin$^-$ BM cells. Percentages obtained with the PGK-*EGFP* reporter donor for the 5' integration junction (in green) or for the 3' integration junction (in blue) calculated in the hematopoietic colonies that were positive for the PCR. WT Lin$^-$ BM cells were nucleofected with different doses of mTALEN (0.75 or 2.5 µg of each monomer) and of the donor (2 or 4 µg). Data are the median ± interquartile range (n=1-8. Each point corresponds to one experiment). No statistical differences were found among groups with a non-parametric Kruskal-Wallis and median test.

Targeted integration frequencies in hematopoietic colonies seeded at 5 days post-nucleofection (from nucleofected cells with 2.5 µg of each mTALEN monomer and 4 µg of the PGK-*EGFP* reporter donor) were also determined. For the 5' integration junction the integration efficiency corresponding to the mean ± S.D. was 8.91 ± 7.34%, and for the 3' integration junction the efficiency was 25.57 ± 19.73%. It is important to highlight that in two out of three of these experiments, 7% and 2% of the analysed colonies were positive for both junctions. The TI frequencies obtained were in the same range as TI frequencies determined when the cells were maintained for 2 days in culture after nucleofection (previous experiments, **Figure 31**). In some exceptions, the specific integration bands also appeared in the samples of cells nucleofected with only 4 µg of donor.

Since it has been recently published that small molecules such as L755507 or BrefeldinA could increase precise genome editing without toxicity [584], we performed one experiment with these molecules. Lin$^-$ BM cells were nucleofected with 2.5 µg of each mTALEN monomer and 4 µg of the PGK-*EGFP* reporter donor, and L755507 (5 µM) and BrefeldinA (0.1 µM) were added immediately after nucleofection. In T+D condition, both the viability (mean value 6.32%) and the transfection efficiency (mean value 25.47%) were similar to that observed in the previous experiments (**Figure 27 A**). Nucleofected Lin$^-$ BM cells treated with BrefeldinA gave rise to a mean TI percentage of 8.83% and 11.76% for the 5' and the 3' integration junctions, respectively. In nucleofected cells treated with L755507, TI values were 7.89% and 0% for the 5' and the 3' integration junctions, respectively. Untreated nucleofected Lin$^-$ BM cells generated colonies with the highest TI percentage, 13.16% and 10.53% for the 5' and the 3' integration junctions, respectively. In these experiments, 2% of TI frequency calculated from colonies positive for both junctions was obtained regardless of whether nucleofected cells were treated with BrefeldinA or not, indicating that none of the tested molecules mediated evident improvements in the TI frequencies of mHSPCs.

Figure 32 and **Figure 33** show representative PCR analyses of the 5' and the 3' integration junctions in hematopoietic colonies.

The results shown in this section constitute the first demonstration of gene editing into the *Mbs85* locus of mouse hematopoietic progenitor cells. The use of 2.5 µg of each mTALEN monomer plus 4 µg of donor gave the highest percentages of targeted integration.

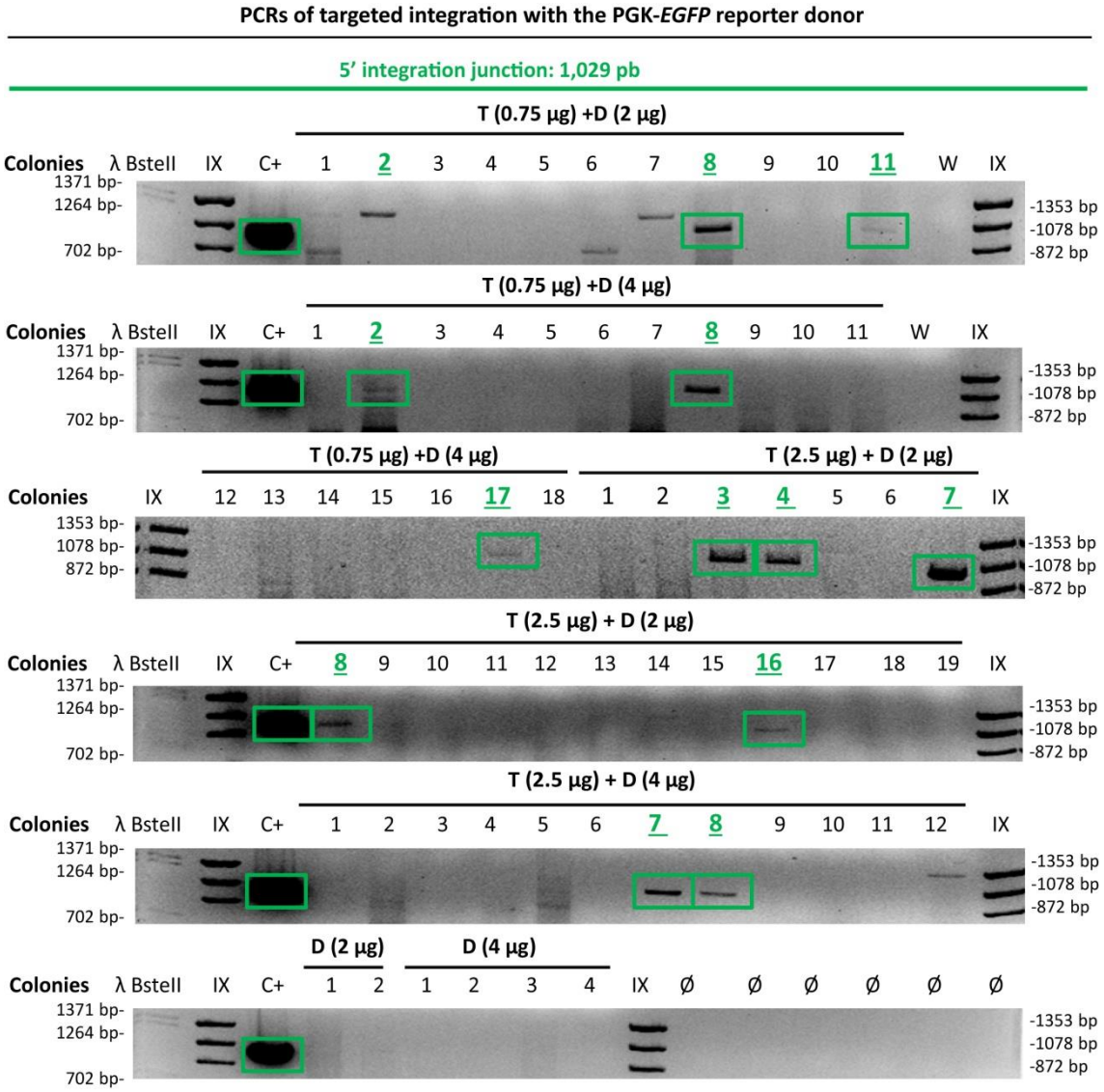


Figure 32. Targeted integration analysis of the PGK-*EGFP* reporter donor into the *Mbs85* locus in colonies derived from gene-edited WT Lin⁻ BM cells for the 5' integration junction. Representative PCR analysis of the 5' integration junction (1,029 pb). ϕ: no sample; W: water; C+: plasmid DNA positive for the 5' integration junction. Samples from nucleofected WT Lin⁻ BM cells with different doses of mTALEN (0.75 or 2.5 µg of each monomer) and of the donor (2 or 4 µg), or with only different donor doses (2 or 4 µg). IX and λ BstII: DNA molecular weight markers. Analysed colonies are numbered and the positive ones framed in green.

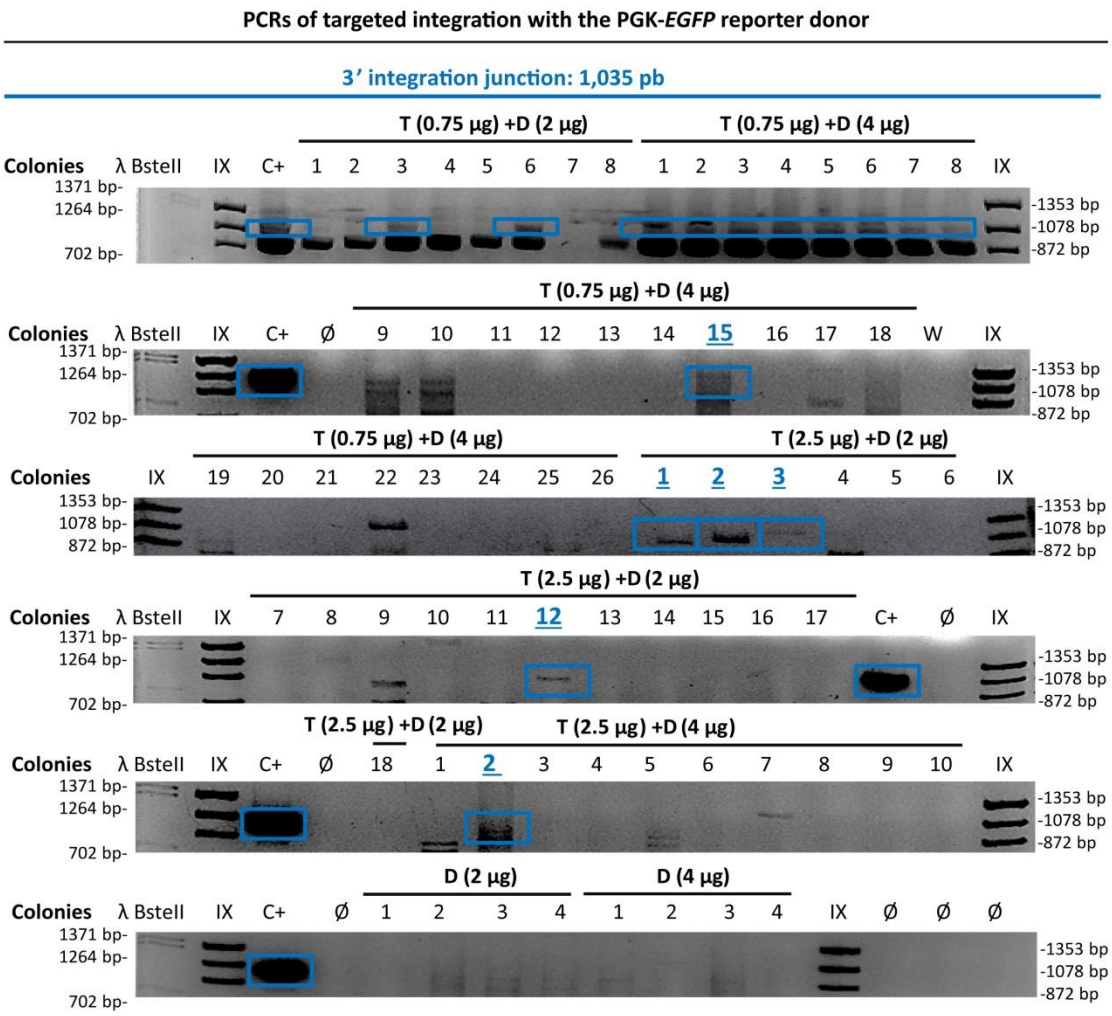


Figure 33. Targeted integration analysis of the PGK-*EGFP* reporter donor into the *Mbs85* locus in colonies derived from gene-edited WT Lin$^-$ BM cells for the 3' integration junction. Representative PCR analysis of the 3' integration junction (1,035 pb). ϕ: no sample; W: water; C+: plasmid DNA positive for the 3' integration junction. Samples from nucleofected WT Lin$^-$ BM cells with different doses of mTALEN (0.75 or 2.5 µg of each monomer) and of the donor (2 or 4 µg), or with only different donor doses (2 or 4 µg). IX and λ BstII: DNA molecular weight markers. Analysed colonies are numbered and the positive ones framed in blue.

3.3.3. Analysis of the *in vivo* repopulating ability of gene-edited WT HSPCs

With the aim of evaluating the ability of WT Lin$^-$ BM gene-edited cells to engraft in recipient mice, a model of congenic transplant of mHSPCs from C57BL/6 mice (CD45.2$^+$) into recipient P3B mice (CD45.1$^+$) was used (details explained in Materials & Methods section 5.3.2).

Cells from experiments Lin$^-$ 29, Lin$^-$ 30, Lin$^-$ 32 and Lin$^-$ 34 were used to perform these studies. In this particular set of experiments, mean TI values in WT Lin$^-$ BM gene-edited cells were calculated. When 0.75 µg of each mTALEN monomer and 4 µg of donor were nucleofected, TI frequencies corresponding to the mean were 17.5% and 7.5% as calculated from in-out PCRs with the 5' and the 3' primers, respectively. When 2.5 µg of each mTALEN monomer and 4 µg of donor were used, the values corresponding to the mean ± S.D. were

9.53 ± 6.41% and 34.03 ± 19.11% for the 5' and the 3' integration junctions, respectively. Most of the nucleofected cells were transplanted into myeloablated mice at 48 hours post-nucleofection. The number of transplanted nucleofected Lin^- BM cells ranged from 20,000 to 626,000 cells. The engraftment values were monitored at 30, 60, and 90 days post-transplant (dpt) by measuring by flow cytometry $CD45.2^+$ cells in the PB of transplanted mice.

As shown in **Figure 34 A**, mock transfected cells (in grey) efficiently engrafted in recipient mice (mean values of 40.57 ± 22.92% at 90 dpt). However, mice transplanted with nucleofected cells either with the mTALEN and the donor (in dark or light green, depending on the dose of mTALEN), or with only the donor (in dark blue) presented very low levels of engraftment at 90 dpt (in most cases below 5%). As expected, mice transplanted with higher doses of nucleofected cells showed higher percentages of engraftment (**Table S5** summarizes the number of cells and transplanted mice in each experiment). Additionally, no EGFP expression was observed by FACs in the PB of these mice.

Multilineage differentiation analyses were performed in the PB of these mice at 30 and 60 days post-transplant. These analyses showed that most recipient mice presented both lymphoid (B220-CD3 markers) and myeloid populations (Gr1-Mac-1 markers). Moreover, this distribution was similar among the different groups of transplanted mice (**Figure 34 B**).

At 90 days post-transplant, donor chimerism was also analysed in the BM of transplanted mice with nucleofected cells (T2.5+D or T0.75+D). Most of the mice presented very low levels of engraftment (**Figure 34 C**). Neither in BM cells nor in the hematopoietic colonies generated from these cells, was EGFP expression observed. Finally, a cell sorting experiment was performed to purify BM $CD45.2^+$ donor cells (details explained in Materials & Methods section 4.1.2). These cells were seeded for clonogenic analysis, but no colonies were generated.

Moreover, neither extending the culture of nucleofected cells from 2 to 5 days after nucleofection, nor the addition of L755507 or BrefeldinA molecules immediately after nucleofection increased the engraftment of the gene-edited cells ($CD45.2^+$) in P3B recipient mice ($CD45.1^+$).

These results indicate that the toxicity associated to the nucleofection of WT Lin^- BM cells with the mTALEN and the donor as plasmid DNA markedly reduced the *in vivo* repopulating properties of these mHSPCs.

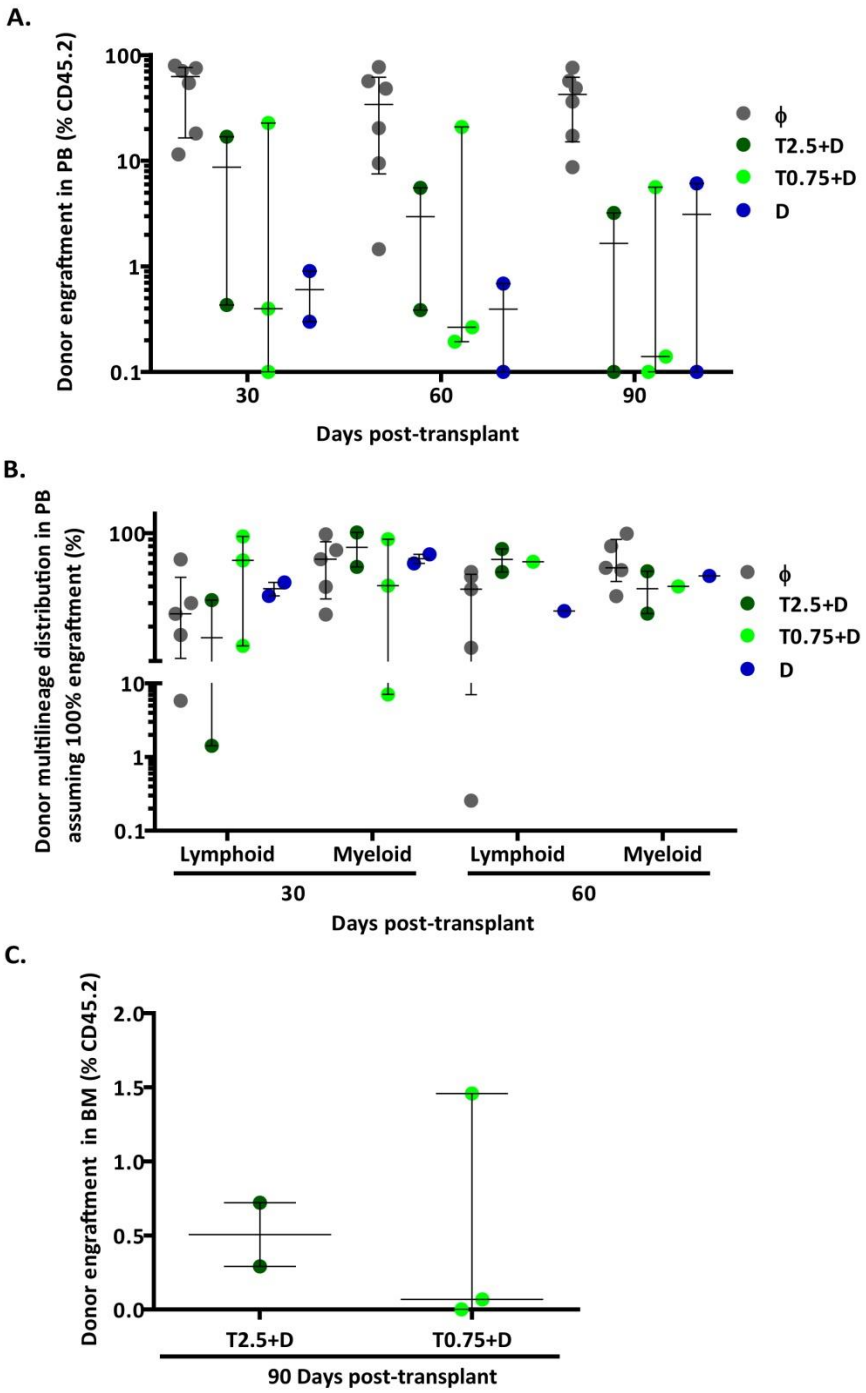


Figure 34. Analysis of chimerism of P3B recipient mice (CD45.1$^+$) transplanted with nucleofected WT Lin$^-$ cells (CD45.2$^+$). **A)** Engraftment measured as the percentage of CD45.2$^+$ cells analysed in PB of P3B recipient mice (CD45.1$^+$) at 30, 60, and 90 dpt. Data are the median ± interquartile range (n=2, 3, or 6 mice per condition from 4 independent experiments). No statistical differences were found among groups with a non-parametric Kruskal-Wallis and median test. **B)** Distribution of donor hematopoietic engraftment of nucleofected cells was analysed by flow cytometry using antibodies: anti-B220-CD3 for B and T cells (lymphoid population) or anti Gr1-Mac-1 for granulocytes and macrophages (myeloid population) at 30 and 60 dpt. Data are the median ± interquartile range (n=2, 3, or 6 mice per condition from 4 independent experiments). No statistical differences were found among groups with a non-parametric Kruskal-Wallis and median test. **C)** Engraftment measured as the percentage of CD45.2$^+$ cells in the BM of P3B recipient mice (CD45.1$^+$) at 90 dpt. Data are the median ± interquartile range (n=2, or 3 mice per condition from 4 independent experiments). Statistical analysis could not be performed. Data referred to are from experiments Lin$^-$ 27, Lin$^-$ 30, Lin$^-$ 32 and Lin$^-$ 34 (**Table S4**).

3.4. Evaluation of mTALEN cleavage efficacy in the *Mbs85* locus of FA-A mouse HSPCs

As in WT mHSPCs, the cleavage efficacy of the NN-mTALEN pair in the *Mbs85* locus of mHSPCs from FA-A mice (FVB strain) was evaluated by Surveyor assay. The cleavage efficacies of the NN-mTALEN pair observed in four independent experiments were 20.46 ± 10.09 % in FA-A Lin$^-$ BM cells (**Figure 35).** These results indicate that although mTALEN had a tendency to cleave in FA-A Lin$^-$ BM cells with a lower efficacy than in WT Lin$^-$ BM cells, no significant differences were found in the cleavage of these cells from different mouse genotypes (**Figure 26 B**).

Therefore, the NN-mTALEN pair cleaved the *Mbs85* locus with similar efficacies in mHSPCs both from WT and FA-A mice.

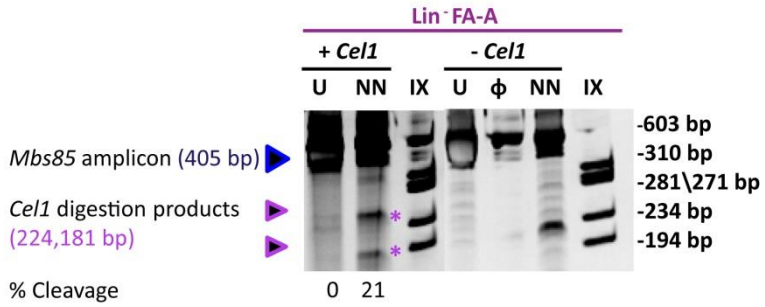


Figure 35. Evaluation of the cleavage efficacy of mTALEN in the *Mbs85* locus of FA-A Lin$^-$ BM cells. Representative electrophoresis gel representing the disruption of the target locus using 2.5 µg of each NN-mTALEN monomer indicated as NN. Samples not digested with the endonuclease were used as controls. The extent of cleavage measured as the mean percentage of modified alleles is indicated. Arrows indicate the size of the parental band (405 bp, in blue) and the expected positions of the digestion products (224 bp and 181 bp, in purple), that are also indicated with purple asterisks. U: untransfected cells; ø: no sample; IX: DNA molecular weight marker.

3.5. Gene-targeting analyses in FA-A hematopoietic progenitors with the therapeutic PGK-h*FANCA* donor

Due to the expected proliferative advantage of corrected HSPCs over the non-corrected ones, Fanconi anemia has been proposed as a relevant disease model to implement gene-targeting approaches in FA-A HSPCs. This process could facilitate the restoration of the hematopoietic system in FA patients.

Considering that we have demonstrated the feasibility of conducting targeted integration in WT hematopoietic progenitors, we wanted to perform similar experiments in mHSPCs from FA-A mice to correct the phenotype of these cells. We performed gene-editing experiments in these cells using the NN-mTALEN pair together with the therapeutic PGK-h*FANCA* donor.

3.5.1. Evaluation of viability, transfection efficiency and clonogenic ability of nucleofected FA-A Lin$^-$ BM cells

As the viability, transfection efficiency and clonogenic ability of nucleofected WT Lin$^-$ BM cells had previously been determined, we evaluated the same parameters in FA-A Lin$^-$ BM cells. Since higher TI frequencies were obtained in WT Lin$^-$ cells with 2.5 µg of each mTALEN

monomer, we used this dose in FA-A experiments. With respect to PGK-h*FANCA* donor, two doses of 2 and 4 µg of DNA were used. DNA-transfected cells showed a 11-fold decrease in the viability either with respect to untransfected or to mock transfected cells. However, in all instances the viabilities ranged from 2.4% to 4.6%. and no significant differences among the different DNA nucleofected conditions were found (**Figure 36 A**). The transfection efficiency obtained with an *EGFP* control plasmid was 28.6%.

When comparing the viability and transfection efficiencies obtained in nucleofected FA-A Lin^- BM cells (**Figure 36 A**) with those obtained in nucleofected WT cells (**Figure 27 A**), no significant differences were obtained among the different conditions (with a two-way ANOVA for the viability and with an one-way ANOVA for the transfection efficiencies). However, we observed a trend suggesting that DNA nucleofected WT Lin^- BM cells (**Figure 27 A**) had higher viabilities in comparison with nucleofected FA-A Lin^- BM cells (**Figure 36 A**).

Regarding the clonogenic assays performed in FA-A Lin^- BM cells, our results showed that DNA nucleofection reduced the clonogenic ability of FA-A Lin^- BM cells approximately 9-fold in comparison with cells subjected to nucleofection in the absence of DNA (**Figure 36 B**). Additionally, no significant differences were observed using a two-way ANOVA test in terms of the clonogenic potential of DNA nucleofected FA-A mHSPCs when compared to the data obtained in DNA nucleofected WT mHSPCs (**Figure 27 C**).

Because the therapeutic cassette contained a puromycin resistance gene, in some experiments 1 µg/ml of puromycin was added to methylcellulose cultures. When 2 µg of the therapeutic PGK-h*FANCA* donor were used, a very low number of colonies was generated regardless of the use of the TALEN. However, puromycin selection showed moderate increases in the number of puro-resistant colonies after gene editing with 2.5 µg of each mTALEN monomer and 4 µg of the donor. The number of colonies generated increased on average 2.6-fold with respect to the samples only nucleofected with the donor (**Figure 36 C**). This difference, however, was not statistically significant.

These experiments show that the viability, transfection efficiency and clonogenic ability of nucleofected FA-A mHSPCs was similar to that observed in WT mHSPCs.

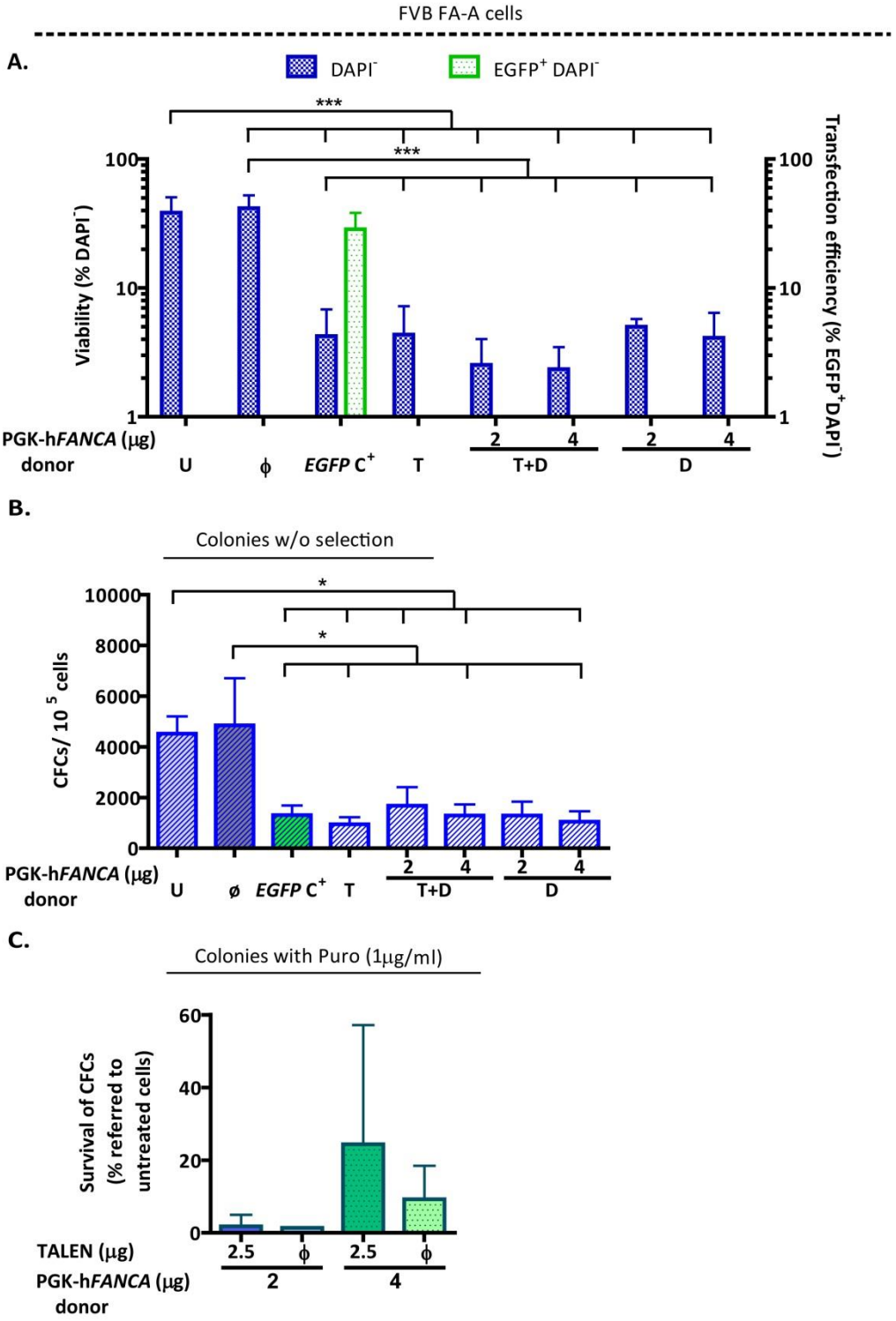


Figure 36. Evaluation of viability, transfection efficiency, and clonogenic ability of nucleofected FA-A Lin⁻ BM cells with the mTALEN and the therapeutic PGK-h*FANCA* donor at 48 hours post-nucleofection. **A)** Analysis of viability (percentage of DAPI⁻ cells) and transfection efficiency (percentage of DAPI⁻ EGFP⁺ cells) determined at 48 hours post-nucleofection. U: untransfected cells; ϕ: transfected cells without DNA (mock condition); *EGFP* C⁺: *EGFP* control plasmid; T: 2.5 μg of each NN-mTALE nuclease monomer; T+D: 2.5 μg of each NN-mTALE nuclease monomer plus the indicated doses of the therapeutic PGK-h*FANCA* donor (2 or 4 μg); D: 2 or 4 μg of the therapeutic PGK-h*FANCA* donor. Data are the mean ± S.D. (n=7 experiments). (***) P-value <0.001 indicates significant differences with a one-way ANOVA followed by a *post-hoc* Tukey test. **B)** Clonogenic assays of nucleofected Lin⁻ cells under the conditions shown in A). Data are the mean ± S.D. (n=7 experiments). (*) P-value <0.05 indicate

significant differences with a one-way ANOVA followed by a *post-hoc* Tukey test. C) Clonogenic assays of the nucleofected Lin$^-$ cells cultured in the presence of 1 µg/ml of puromycin. Percentages with respect to colony numbers obtained without drug selection. ϕ: transfected cells with the donor without mTALEN. Bars show the mean ± S.D. (n=4 experiments, except in D2 µg where n=1). No statistical differences were found among groups with a non-parametric Kruskal-Wallis and median test. Data referred to are from experiments Lin$^-$ 9; Lin$^-$ 10; Lin$^-$ 11; Lin$^-$ 12; Lin$^-$ 21; Lin$^-$ 22; and Lin$^-$ 23 (**Table S4**).

3.5.2. Evaluation of the phenotypic correction of FA-A Lin$^-$ BM cells complemented by gene targeting

To test the phenotypic correction of gene-edited FA-A Lin$^-$ BM cells, cells were plated in methylcellulose in the presence and the absence of different doses of MMC (10 and 30 nM).

With 10 nM of MMC, the survival of hematopoietic colonies derived from samples nucleofected with 2 µg of donor plus the mTALEN did not increase in comparison with that obtained in samples nucleofected with only the donor. However, when the donor dose was increased to 4 µg, a 1.7-fold increase in the survival of hematopoietic colonies was observed in samples nucleofected both with the mTALEN and the donor in comparison with samples nucleofected with only the donor (**Figure 37**).

Moreover, when 30 nM of MMC were used, the survival of hematopoietic colonies derived from samples nucleofected with the mTALEN and 4 µg of donor increased 34.1-fold in comparison with samples nucleofected with only the donor. Specifically, we obtained on average 6% MMC resistant colonies in samples nucleofected with mTALEN and 4 µg of donor, in comparison to 0.2% MMC resistant colonies when only the donor was used (**Figure 37**).

Thus, analyses of resistance to 30 nM of MMC performed in gene-edited FA-A mHSPCs suggested efficacies of phenotypic correction of 6% under optimal conditions.

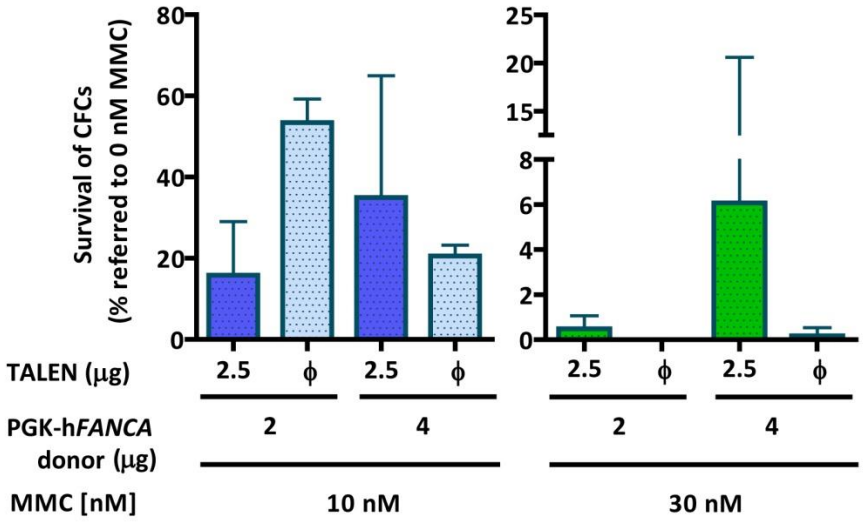


Figure 37. Clonogenic potential of nucleofected FA-A Lin$^-$ BM cells treated with 10 and 30 nM of MMC compared to cells cultured in the absence of MMC. Clonogenic assays were performed to evaluate the ability of nucleofected Lin$^-$ BM cells to generate hematopoietic colonies in the presence of 10 or 30 nM of MMC. MMC survivals are indicated considering the number of hematopoietic colonies generated without drug selection as 100%. Studied conditions: 2.5 µg of each NN-mTALE nuclease monomer were nucleofected except in condition ϕ where cells were nucleofected with only the therapeutic PGK-h*FANCA* donor. Two doses of 2 and 4 µg of the therapeutic PGK-h*FANCA* donor were used for nucleofection as is indicated. Data are the mean ± S.D. (n=1, 2, 4, 5, 7 or 8 experiments). No statistical

differences were found among groups with a non-parametric Kruskal-Wallis and median test. Data referred to are from experiments Lin⁻ 9; Lin⁻ 10; Lin⁻ 11; Lin⁻ 12; Lin⁻ 14; Lin⁻ 15; and Lin⁻ 23 (**Table S4**).

3.5.3. Analysis of targeted integration efficiency of the therapeutic PGK-h*FANCA* donor into the *Mbs85* locus of FA-A Lin⁻ BM cells by PCR

PCR analyses of hematopoietic colonies grown in presence of puromycin or MMC (from three and eight experiments, respectively) were performed to analyse the TI frequency of the therapeutic PGK-h*FANCA* donor into the *Mbs85* locus of FA-A hematopoietic progenitors. In-out PCRs for the 5' and the 3' integration junctions were conducted.

The estimated mean TI frequencies obtained in puromycin resistant hematopoietic colonies were in the range of 3.0% to 8.8%. When 10 nM of MMC was used, the mean TI frequency in the surviving colonies was in the range of 2.9% to 10.1 %, while in the colonies resistant to 30 nM, this range varied between 0% and 16.5%. In each selection method, except in the T+D2 condition with 30 nM of MMC (n=2) where no bands of PCR amplification were detected for the 5' junction, we obtained integration frequencies that were in the same range for both integration junctions (**Figure 38 A**).

The estimated frequencies regardless of the method of selection in the analysed colonies corresponding to the mean ± S.D., using 2 µg of donor were 5.81 ± 8.57% and 9.51 ± 15.33% for the 5' and the 3' integration junctions, respectively. When 4 µg of donor were used, the corresponding values were 7.73 ± 7.49% and 8.15 ± 11.91% for the 5' and the 3' integration junctions, respectively (**Figure 38 B**).

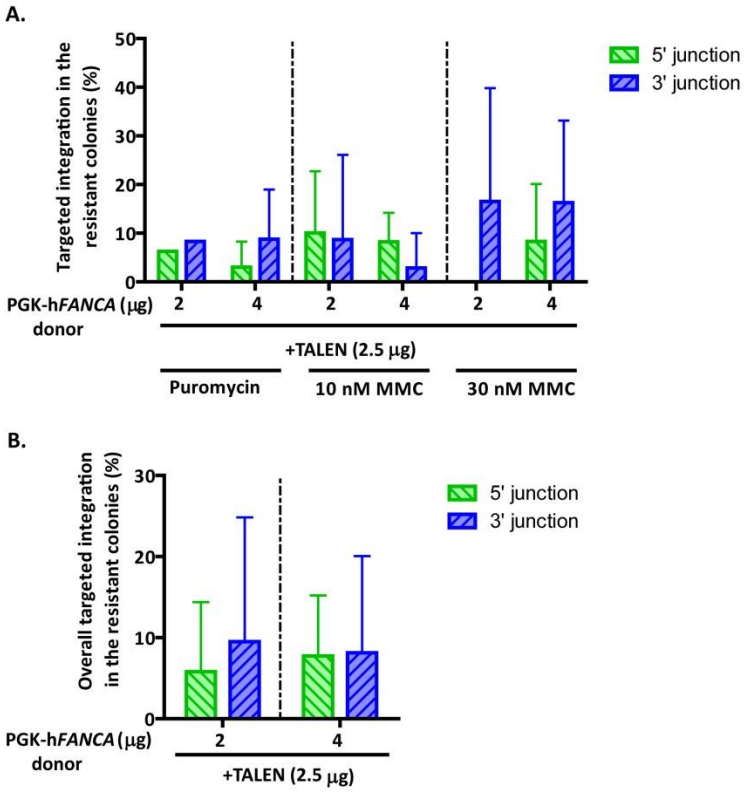


Figure 38. Estimated frequencies of targeted integration of the therapeutic *hFANCA* donor into the *Mbs85* locus of FA-A Lin⁻ BM cells. A) Percentages of targeted integration calculated in PCR analyses for the 5' integration junction (in green) or for the 3' integration junction (in blue) in the different

conditions of T+D (2 µg) or T+D (4 µg). Percentages were calculated in hematopoietic colonies resistant after the treatment with puromycin or MMC (10 or 30 nM). Data are the mean ± S.D. (n=1-8 experiments). No statistical differences were found among groups with a non-parametric Kruskal-Wallis and median test. **B)** Percentages calculated from the PCR positive colonies resistant to puromycin or to MMC (10 or 30 nM), including all selection methods for the 5' integration junction (in green) or for the 3' integration junction (in blue) in the different conditions of T+D (2 µg) or T+D (4 µg). Data are the mean ± S.D. (n=8-14 experiments). No statistical differences were found among groups with a parametric two-way ANOVA test.

Figure 39 and **Figure 40** show two representative PCRs from two different experiments. The first figure corresponds to the analyses of puromycin resistant colonies, and the second one corresponds to the analyses of 10 and 30 nM MMC resistant colonies, in which both the 5' and the 3' integration junctions were analysed. As described in previous sections, in only a few colonies was possible to observe the specific integration bands for both integration junctions. In some experiments however, we were able to amplify both bands in the same colony, the average TI efficiency being 5.04%. As happened in WT samples, we could not detect any PCR amplification either in the 5' or in the 3' integration junctions in colonies derived from samples nucleofected with only the donor.

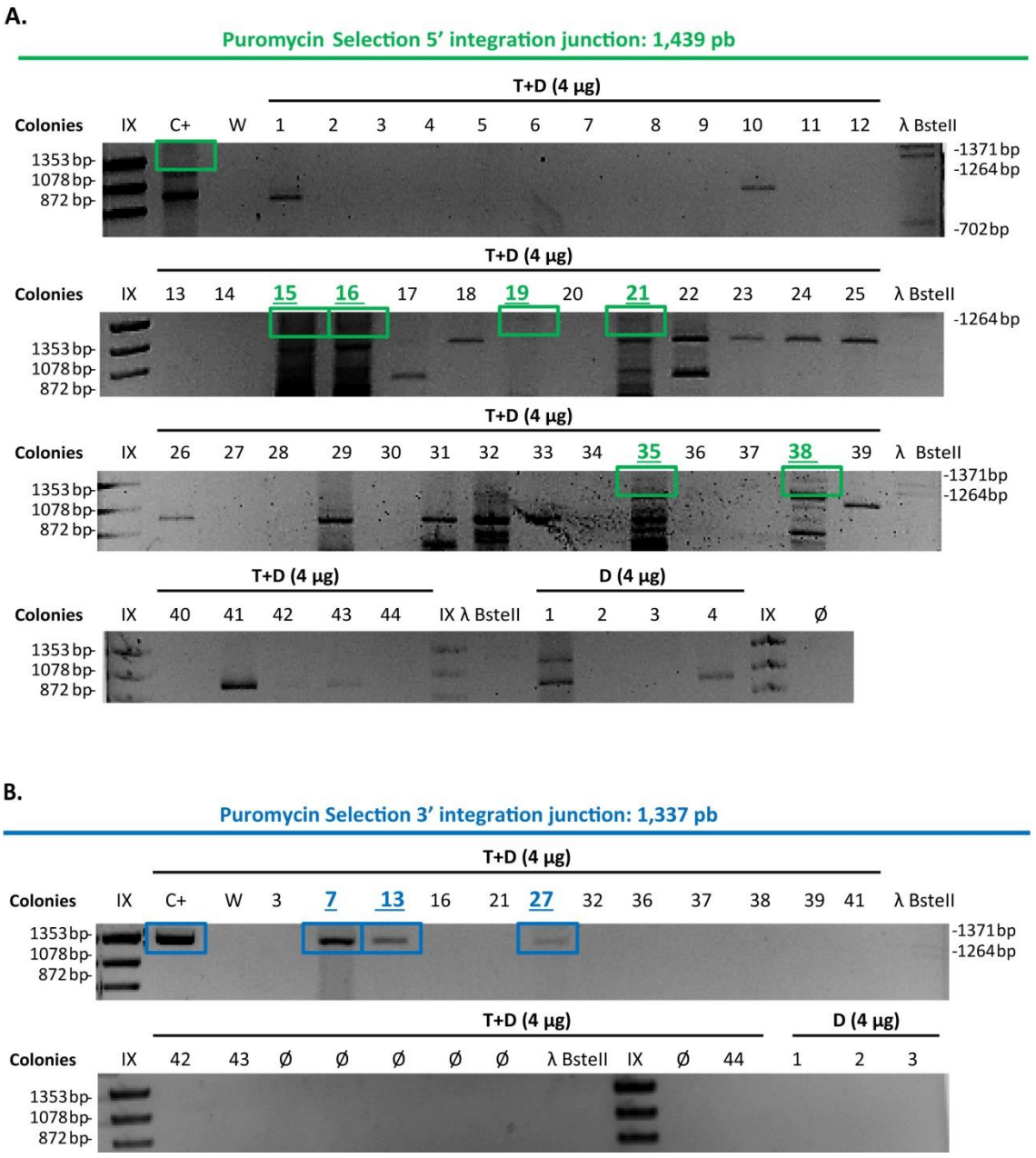


Figure 39. Targeted integration analysis of the therapeutic PGK-*hFANCA* donor in colonies derived from gene-edited FA-A Lin⁻ BM cells resistant to puromycin. Representative PCR analyses of the same experiment for the study of the 5' integration junction (1,439 pb) **(A)** and the 3' integration junction (1,337 pb) **(B)**. ϕ: no sample; W: water; C+: sequenced genomic DNA positive for the 5' or the 3' integration junctions in gene-edited FA-A MEFs; T+D (4 µg): 2.5 µg of each NN-mTALE nuclease monomer plus 4 µg of the therapeutic PGK-h*FANCA* donor; D (4 µg): 4 µg of the therapeutic PGK-h*FANCA* donor; IX and λ BstII: DNA molecular weight markers. Analysed colonies are numbered and the positive ones framed in green (5' integration junction) or in blue (3' integration junction).

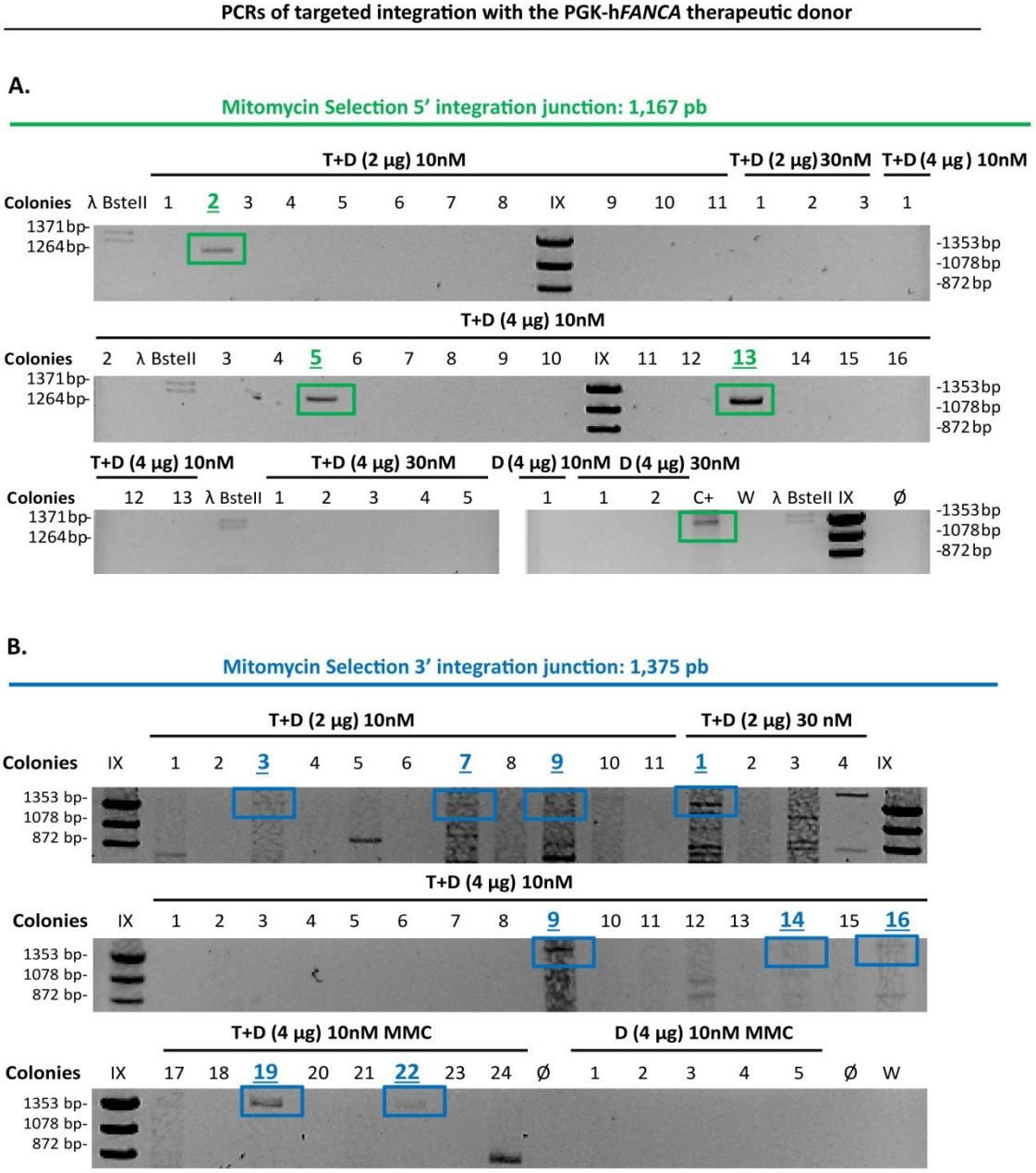


Figure 40. Targeted integration analysis of the therapeutic PGK-*hFANCA* donor in colonies derived from gene-edited FA-A Lin⁻ BM cells resistant to 10 or 30 nM of MMC. Representative PCR analyses of the same experiment for the study of the 5' integration junction (1,167 pb) **(A)** and the 3' integration junction (1,375 pb) **(B)**. ϕ: no sample; W: water; C+: sequenced genomic DNA positive for the 5' or the 3' integration junctions in gene-edited FA-A MEFs; T+D (2 or 4 µg): 2.5 µg of each NN-mTALE nuclease monomer plus 2 or 4 µg of the therapeutic PGK-h*FANCA* donor; D (4 µg): 4 µg of the therapeutic PGK-h*FANCA* donor; IX and λ BstII: DNA molecular weight markers. Analysed colonies are numbered and the positive ones framed in green (5' integration junction) or in blue (3' integration junction).

We analysed the total targeted integration percentages both considering the survival rate of the hematopoietic colonies after selection and also the percentage of colonies that were positive for PCR analyses either for the 5' or the 3' targeted integration junctions (shown in **Figure 41**).

When 2.5 µg of each mTALEN monomer and 2 µg of donor were used, mean TI frequencies in colonies selected with puromycin varied from 0.34% to 0.45% for the 5' and the 3' integration junctions, respectively. When 4 µg of donor were used together with 2.5 µg of each mTALEN monomer, TI frequencies corresponding to the mean ± S.D. were 2.19 ± 3.79% and 5.19 ± 8.03% for the 5' and the 3' integration junctions, respectively. In colonies selected with 10nM of MMC, when cells were nucleofected with 2.5 µg of each mTALEN monomer and 2 µg of donor, TI frequencies were 1.27 ± 1.62% and 0.33 ± 0.74% for the 5' and the 3' integration junctions, respectively. And when the dose of donor was increased up to 4 µg, TI frequencies of 4.25 ± 3.15% and 1.49 ± 3.50% were achieved in the 5' and the 3' integration junction analyses. Finally, in colonies selected with 30 nM of MMC, mean TI frequencies in hematopoietic colonies from samples nucleofected with 2.5 µg of mTALEN and 2 µg of donor were of 0% and 0.14% in the 5' and the 3' integration junctions analyses, respectively. But TI frequencies in hematopoietic colonies from samples nucleofected with 2.5 µg of mTALEN and 4 µg of donor reached 2.72 ± 4.85% and 1.90 ± 2.08% in the 5' and the 3' integration junctions analyses, respectively. Furthermore, we observed that the mean targeted integration percentages in the total population was higher using 4 µg of donor as compared with 2 µg of donor, regardless of the selection method.

Thus, these results demonstrate the phenotypic correction of FA-A mHSPCs through the integration of the therapeutic PGK-h*FANCA* donor into the *Mbs85* locus of these cells.

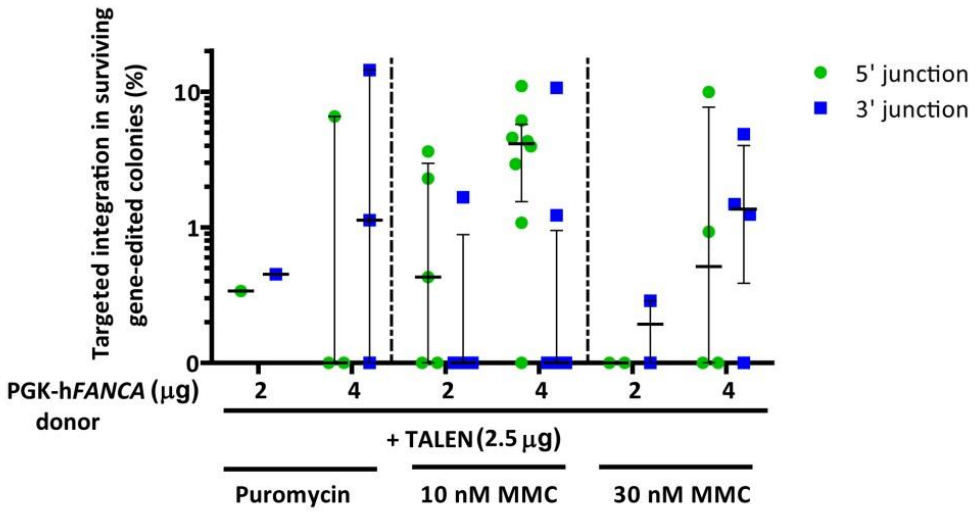


Figure 41. Targeted integration percentage of the therapeutic PGK-*hFANCA* donor into the *Mbs85* locus in surviving colonies derived from gene-edited FA-A Lin$^-$ BM cells. TI percentages achieved with the different selection methods for the 5' integration junction (in green) and for the 3' integration junction (in blue) considering the percentage of surviving hematopoietic colonies and those that were positive for the PCR. Data show the median ± interquartile range (n=1-8. Each point corresponds to one experiment). No statistical differences were found among groups with a non-parametric Kruskal-Wallis and median test.

3.5.4. Analysis of the *in vivo* repopulating ability of gene-edited FA-A HSPCs

As our final goal was to restore the hematological phenotype of FA-A mice using a gene-editing approach, we aimed at analysing if the levels of targeted integration would be enough to repopulate the hematopoiesis of myeloablated FA-A mice.

Three experiments (Lin$^-$ 14, 15 and 17) were performed with this objective in mind. Donor cells were always from male FA-A mice while recipients were always FA-A female mice. Chimerism was analysed by qPCR, calculating the percentage of positive cells for the *Sry* gene in the PB of FA-A recipient mice (as explained in Materials & Methods, section 5.3.1).

In these experiments, from 28,000 to 571,250 nucleofected cells were transplanted into each recipient. Nevertheless, all recipients died between seven and fourteen days post-transplant due to graft failure.

One experiment was then performed to analyse what the minimum dose of untransfected FA-A Lin$^-$ BM cells required to engraft FA-A recipients was. Three doses of untransfected cells were transplanted: $3x10^5$, $4.5x10^5$ and $6x10^5$ Lin$^-$ BM cells in myeloablated mice (**Figure 42**). At 120 days post-transplant only a few of them showed donor engraftment (1 out of 4 mice in the groups transplanted with $3x10^5$ and $6x10^5$ cells, and 2 out of 4 mice in group transplanted with $4.5x10^5$ cells).

Taken together, this section of results demonstrate for the first time a therapeutic targeted gene integration into the *Mbs85 safe harbor* locus of progenitor cells from a mouse model of FA-A. As happened with WT mHSPCs, the toxicity associated with the nucleofection of these cells with the TALEN and the donor as DNA constructs restricted the repopulating potential of the primitive mHSPCs.

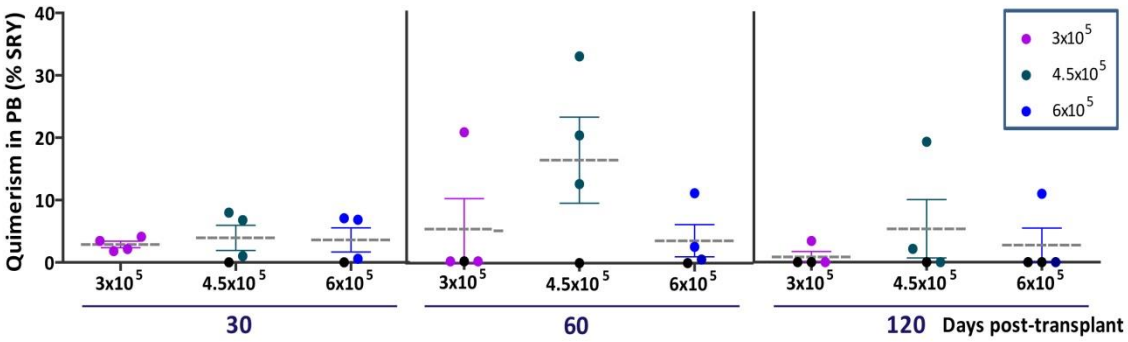


Figure 42. Analysis of chimerism by qPCR of the untransfected FA-A Lin$^-$ BM cells transplanted in FA-A recipients. Hematopoietic graft percentages with different doses of untransfected male FA-A Lin$^-$ BM cells ($3x10^5$ in purple, $4.5x10^5$ in green, and $6x10^5$ cells in blue) analysed by qPCR in the PB of FA-A female recipients at 30, 60 and 120 dpt. Mice that died are colored in black. Data referred to are from experiment Lin$^-$ 19 (**Table S4**).

IX. DISCUSSION

DISCUSSION

As it has already mentioned in the introduction, targeted gene-editing strategies based on HDR would circumvent some of the limitations of conventional GT, mainly the impossibility of controlling the insertions in the genome. In this respect, the main parameters to take into account in a gene-targeting strategy are the efficiency, the specificity and the fidelity of the process [327].

In this work, we have investigated the use of TALEN to promote the specific insertion of either a *EGFP* reporter cassette or of a therapeutic cassette carrying the human *FANCA* gene into a potential *safe harbor* locus, the *Mbs85* locus of WT or FA-A cells. The ultimate goal of this project was to achieve engraftment of the gene-edited FA-A mHSPCs in FA-A mice, with the aim of correcting their characteristic hematopoietic phenotype.

1. EFFICIENCY AND SPECIFICITY OF THE DESIGNED *Mbs85* TALEN

Up to now, TALEN have been described as a very specific platform to generate DSBs due to their long DNA binding sequence and also because they function as an obligate heterodimer [328]. Although ZFN are the only nucleases that have been used in the clinic so far, probably because they were the first ones to be developed, TALEN were chosen for our strategy because of their specificity that would limit their number of potential off-targets and related cytotoxicity [376, 378]. The first experiments consisted in the analysis of the efficiency of our mTALE nuclease to cleave the *Mbs85* locus. Additionally, we have also analysed the potential off-targets of these TALEN. The cleavage capacity in the on-target locus without generating off-targets will define the specificity of the mTALENs designed to target the *Mbs85* locus.

First, we validated the expression of the different pairs of mTALEN (differing in the repeat-variable di-residues (RVD) in the targeting modules that recognize G nucleotides) after their lipofection as plasmid DNA in HEK-293T cells. HEK-293T cells are characterized by their easy handling, high rate of proliferation and easiness to be transfected. These particular features make them a very suitable cell line to test the expression of the different pairs of nucleases. Our data in **Figure 14 A** showed that the viability was not highly compromised (mean value of 59% $DAPI^-$ cells) and the transfection efficiencies were significant (36% $DAPI^- EGFP^+$ cells). WB analysis showed the expression of the HA-tag domain of the TALEN backbone from all mTALEN monomers that contained different targeting modules with different specificities (**Figure 14 B**).

Once we analysed the mTALEN expression, we validated their efficacy in introducing DSBs in the target *Mbs85* locus. INDELs generated after NHEJ repair provided an indirect measurement of the cleavage efficacy of the engineered nucleases. We analysed two pairs of mTALEN in different cell types: NIH/3T3 cells, Ba/F3 cells and FA-A MEFs (**Figure 15**). Ba/F3 cells were chosen because these cells grow in suspension, similarly to our final cell type (Lin^- BM mHSPCs). FA-A MEFs were chosen as a primary cell line derived from *Fanca*$^{-/-}$ mice in which

we wanted to perform our gene-editing strategy as a proof of concept. In general, it has been established than an active nuclease should have greater than 10% INDEL activity, hence, the mTALEN should have an on-target activity above this threshold. Moreover, the initial studies performed with TALEN containing different targeting modules with different amino acids, (resulting in different specificities) pointed out that the NN-TALENs presented a base preference either to guanine (G) or to a second base, adenine (A), whereas NK-TALENs showed a much stronger preference for G nucleotides [350]. Based on this potential advantage for the recognition of the G base, higher activities were expected from NK-TALENs recognizing G. However, most recent studies demonstrated that NN-TALENs had a higher affinity for targets with G than their NK-counterparts [77, 104].

Considering all these studies, we analysed which pair of nucleases was more active. We observed that the NN-mTALEN pair was more active in Ba/F3 and FA-A MEFs in comparison with the NK-mTALEN pair. Additionally, in these studies, the percentage of cleavage with this pair exceeded the threshold of 10% cleavage efficacy in all the cell types analysed (**Figure 15**). These results were in accordance with the aforementioned studies performed with different NN-pairs of TALENs. Nevertheless, the transfection efficiency or the cell type could also have modulated the cleavage efficacy of the different mTALEN pairs. With respect to cell dose, 2.5 µg of each NN-mTALEN monomer gave the highest percentages of cleavage in primary FA-A MEFs. Therefore, the NN-mTALEN pair used at a dose of 2.5 µg was chosen for our gene-editing strategy in the mouse *Mbs85* locus of FA-A MEFs.

Regardless of the nuclease technology, it is difficult to determine the full spectrum of off-target cleavage that would define the specificity of the TALEN. Different methods to detect potential off-targets have been developed (Sanger sequencing of individual clones, restriction fragment length polymorphism (RFLP) analysis, mismatch-sensitive enzymes), but high-throughput sequencing of bulk populations of cells is the preferred one to detect INDELs induced at low frequencies. Other methods for measuring off-targets in an unbiased fashion and in a sensitive way have been recently reviewed [85, 213, 275]. In this thesis, we have perform an *in silico* identification of the potential 48 top off-targets of the NN-mTALEN pair, scanning the genome with the PROGNOS software that searched related sequences to the *Mbs85* on-target site (**Table 9**). In the near future, we will individually analyse at least the top 20-25 potential off-targets in a bulk cell population of both WT and FA-A MEFs, and Lin^- BM cells by PCR amplification and high-throughput sequencing.

Taken together this first section of results, we have demonstrated the efficient cleavage of the mTALE nucleases in the *Mbs85* locus of different cell types, including primary FA-A MEFs.

2. GENE-TARGETING STUDIES IN FA-A MEFs

As the FA/BRCA pathway is involved in the regulation of HRR, we wanted to evaluate the feasibility of implementing for the first time a gene-targeting strategy in cells from a mouse model of Fanconi anemia corresponding to complementation group A.

Up to now, it has been described that mutations in proteins of the FA core complex result in mild HRR defects in mammalian cells [381], except from replication-dependent repair of the ICLs [382]. Furthermore, it has been described that although FA proteins of the core complex promote HRR, they are not essential components of this pathway [381]. However, FANCA and BRCA1 constitutively interact with each other [159]. Additionally, it has been described that FA-A deficiency in MEFs could enhance the error-prone NHEJ repair [1, 53]. Thus, mutations in *FANCA* could affect the efficiency or activation of the HRR as the preferred pathway to repair DSBs. Notably, our group has already demonstrated the effectiveness of a GT approach nuclease-mediated in fibroblasts from FA-A patients [456]. Consequently, taking into account all these studies, we wanted to implement a gene-editing strategy in mouse *Fanca*$^{-/-}$ cells, with the aim of demonstrating for the first time the feasibility of conducting a therapeutic gene-editing approach first in FA-A MEFs and then in hematopoietic progenitors.

Moreover, taking into account the wide spectrum of mutations in FA, gene editing by HDR in a *safe harbor* locus would be an exceptional pre-clinical approach for future therapeutic gene-editing strategies for all FA mutations and subtypes.

Once we established that the NN-mTALEN pair was the most efficient in cleaving the *Mbs85* locus, we proceeded to perform gene-editing experiments in FA-A MEFs. With this aim in mind, a donor construct containing the therapeutic h*FANCA* gene was generated. The therapeutic h*FANCA* cassette was designed to be expressed under the regulation of the human PGK promoter. This promoter was chosen because it has already been shown to be effective and safe in clinical studies [32], including the clinical trial of FA patients. In addition, a 3XFLAG domain was included in the h*FANCA* cassette to evaluate the transgene expression either by WB or by FACS analyses. Finally, the length of the HAs has been proposed to have an effect on homologous recombination (HR) frequency. It has been proved that HR increases exponentially with the size of the homologous flanking region up to 1 kb [282], although in other studies, lower HAs of 0.3 Kb have been shown to be equally efficient in promoting HRR [425]. In our case, the left and right HAs present in our donor cassettes were 806 bp and 860 bp in length, respectively.

Additionally, we considered different delivery methods both for the TALEN and the donor. At the time we started our study, the preferred approach for donor delivery was based on plasmid DNA. Other strategies, such as the generation of IDLV vectors with high viral titers (to efficiently transduced HSPCs, our main targets cells), are difficult when the donor template is higher than 5 Kb, as it was in our case.

With respect to the TALEN delivery, TALEN should be provided in a “hit-and-run” process, allowing short-term and dose-controllable nuclease activity. Up to now, TALEN

delivery has remained the major obstacle in many studies. Unlike other nuclease platforms, the long and unstable nature of the tandem repeats of TALEN has prevented their packaging in LVs [220]. Moreover, the use LVs with mutated reverse transcriptase for mRNA delivery of TALEN has also been used, but translation levels were not high enough to facilitate the gene editing of primary T cells [356]. AdVs that do not integrate in the genome have been proved to be a valuable delivery platform of artificial nucleases, but their efficacy in transducing HSPCs has classically been very limited [297]. Nevertheless, the use of helper-dependent Ad5/35 expressing ZFN under the regulation of a microRNA in HSCs has recently been published [477]. Other approaches have been studied recently, one of them based on the use of cell-penetrating peptides. However, depending on the locus and ratio of peptide: protein, different levels of mutagenesis were obtained that ranged from very low (<4%) to moderate levels (~16%) of disruption, indicating that further improvements would be necessary [310]. Another approach with LVs as carriers of designed nuclease proteins has been tested. In that study, the authors obtained very low targeted disruption frequencies (<6%), and TI frequencies with a donor template of ~0.2%. Additionally, the authors observed proteolytic cleavage within the TALEN domain of the fusion protein that led to truncated versions of the protein. Therefore, the authors claimed that further studies must be carried out to improve the delivery of virus-incorporated TALEN proteins [54]. Nevertheless, several groups have demonstrated the use of mRNA to successfully deliver nucleases in human $CD34^+$ cells [184, 317, 551]. In this respect, TALEN have been also used as mRNA to efficiently disrupt the *CCR5* locus (being >50% in primary T cells) [357] or the *BCL11A* enhancer (being >40% in $CD34^+$) [551]. Moreover, mRNA delivery has been proposed to offer a rapid burst of expression of TALEN with modest toxicity.

In this study, we investigated mTALEN delivery both as a plasmid DNA and mRNA. However, when we delivered the mTALEN as mRNA, we did not succeed in introducing them in the target cells. Despite the fact that further analysis would be required, we thought that the mTALEN generated as mRNA were not very stable due to their large repeat domains and ease to be degraded during transfection. Thus, since our data with mTALEN as mRNA did not improve the results obtained with plasmid DNA in FA-A MEFs, we continued our experiments with the mTALEN as plasmid DNA.

To facilitate the gene editing mediated by TALEN in FA-A MEFs, these cells were immortalized by transduction of γ-RVs encoding HPV *16 E6-E7* genes. HPV oncoproteins function by deregulating cell cycle checkpoints, or evading replicative senescence caused by telomere shortening, by upregulating hTERT expression or independently of hTERT, by promoting the alternative lengthening of telomeres (ALT) that requires HRR machinery [505]. In our experiment, the immortalization process maintained the hypersensitive phenotype of FA-A cells to MMC, in comparison with their immortalized WT counterparts (**Figure 16 B**). Rio et al. [456] published that the immortalization of human FA-A fibroblasts with h*TERT* increased the percentage of gene targeting in comparison with non-immortalized fibroblasts. In our case, however, we could not make this comparison because non-immortalized mouse embryonic fibroblasts rapidly developed senescence in culture.

Immortalized FA-A MEFs were nucleofected with both the mTALEN and the therapeutic h*FANCA* cassette. In nucleofected cells, the transfection efficiency was reasonably effective (mean value of 30%) and the viabilities were very high (they ranged from 47% to 78%),

indicating that in this cell type, cytotoxicity was not a limiting factor to perform gene editing. We observed, however, a trend indicating that higher donor doses decreased the viability of these cells (**Figure 17 A**). In immortalized FA-A MEFs, mTALENs were very effective because they have a high percentage of cleavage efficacy (mean value of 22.61%) (**Figure 17 B)**. In these experiments we also observed that when a donor template was used together with the mTALEN, the percentage of repair by NHEJ was reduced. This difference in the cleavage could be due to either to the transfection efficiency, or due to a decrease in the number of mismatches generated by NHEJ as HRR is more favoured in the presence of a donor template.

During the repair of the DSBs generated by nucleases, different outcomes could be obtained [327]. The fidelity of the process would be a result of the manner in which the exogenous DNA is inserted at the target site, without the generation of adverse genome-modifying events. In this respect, both the specificity of the nuclease and the nature and structure of the DNA donor component may play a key role. The integration of the donor could take place in the on-target locus, but also where the nuclease has potentially generated off-targets. Additionally, it could be integrated in other breakage-prone fragile sites intrinsic to the cell type or to the disease cells, as well as in other DSBs created by DNA metabolic processes, environmental mutagens or other factors. In our case, these factors need to be considered because we were working with disease cells in which a DNA repair pathway is affected. In addition, concatemeric sequence insertions generated through NHEJ may also been generated. In our study, as we have worked with plasmid DNA donors, other sequences derived from the backbone of the plasmid, such as viral or bacterial DNA, could also be introduced. Plasmid and viral DNA, have been thought to be more prone to random integration events [167, 218, 442]. However, alternatives such as single strand oligonucleotides (ssODNs) have also been shown to present some limitations, particularly, the presence of low fidelity integrants with unintended mutations [92, 439].

Therefore, once we nucleofected the bulk of FA-A MEFs, we analysed the on-target integration of the therapeutic cassette. The approach that we have followed consisted of performing PCRs with primers that hybridized both inside the donor and outside the HAs at the 5' and the 3' integration junctions (in-out PCRs). With this procedure we could not rule out concatemeric or random integration events, but it was useful to analyse whether on-target integrations had taken place. Although the therapeutic h*FANCA* cassette was a large donor of 10,303 bp, and despite the fact that FA-A cells could be defective in HRR [381], we confirmed specific bands in the in-out PCRs when both the donor and the mTALEN were nucleofected together (confirmed in the conditions of T+D0.75 µg and T+D4 µg). On the contrary, when these cells were nucleofected with only the donor, there was no amplification of the specific band. These results suggest the generation of HRR-mediated integrations of the therapeutic donor at the on-target locus (**Figure 18 B**). However, the occurrence of off-target integration events was also observed in our FISH studies (**Figure 20**).

In subsequent experiments, we generated different clones by limiting dilution to calculate the frequencies of gene editing in puromycin-selected FA-A MEFs. PCRs were conducted for amplification of the 5' and in the 3' integration junctions and the observed frequencies based on the clones that were positive for both junctions were of 1.96% and 7.35% in two different experiments (**Table 10**). Significantly, we confirmed by Sanger

sequencing the band amplified by PCR for the 3' integration junction in the bulk of gene-edited FA-A MEFs and in clone #76# (**Figure 18 D**). Strikingly, the 5' integration junction was in many cases difficult to detect, and this could be either due to the efficiency of the PCR, to the genomic instability of these cells, or to the generation of on-target integrants, that for the 5' integration junction could have been improperly repaired. These types of events have been previously described in different cell types [184, 317]. In a recent paper the authors measured the knock-in efficiencies of an *EGFP* donor in different cell types mediated by HDR (ranging from 0.07% to 4.98%) or by NHEJ approaches (ranging from 0.8% to 14.33%). In this study, the authors stated that the majority of the DSBs were repaired by the DNA-PKcs dependent NHEJ (D-NHEJ) pathway [207]. In our particular case, taking into account that FA-A MEFs had been prestimulated for 48 hours before the nucleofection, we hypothesized that the backup-NHEJ (B-NHEJ) or the SSA pathways could have competed with the HDR, which have been described as being more active during S/G2 phases of cell cycle [355].

Comparing our gene-targeting efficiencies (1.96% and 7.35%) with other published studies in monogenic diseases in which the authors also used nucleases and the donor as plasmid DNA, we observed that, Menon et al. using TALEN obtained a 2.6% TI frequency in iPSCs derived from a SCID-X1 patient [345]. In the study of Shamin et al., the authors used ZFN and obtained a TI frequency of 30% in mouse radiosensitive SCID fibroblasts selected with G418; and the targeted frequencies were of 89% in the iPSCs derived from SCID fibroblasts, and extended PCR analyses were only performed in 5 out of 46 clones in which TI was confirmed for both junctions [443]. Interestingly, there are two publications in which the authors developed a gene-targeting strategy in human FA-A or FA-C fibroblasts. In the first work, Rio et al. [456] used a donor IDLV and AdV5/35-ZFNs to target the human *AAVS1* locus. In this study, initial TI frequencies were between 0.2% and 1.1%, however after 42 days, a significant expansion of gene-edited cells was observed (up to 13.4-40%). In the second work, Osborn et al. [403] using the Cas9 as a nuclease, they obtained a TI frequency of 9% in puromycin-selected clones. This percentage increased using the Cas9 as nickase. Additionally, in a recent study, Osborn et al. [402] showed HDR efficiencies of 66-70% in iPSCs derived from FA-I fibroblasts, however, they were not able to generate clones from the bulk of puromycin-selected fibroblasts corrected by HDR. Thus, although none of these studies are exactly comparable, our efficiencies can be considered to be in the same range as those observed in the above-mentioned studies.

The specific integration of the PGK-h*FANCA* cassette into the *Mbs85* locus was also confirmed by FISH. With this technique we could confirm TI of the therapeutic cassette into chromosome 7 (where the *Mbs85* locus is located) in clones #7# and #76# (**Figure 20 A** and **B**). However, we also observed random integration events of the PGK-h*FANCA* cassette in other chromosomes in some metaphase spreads of clones #76#, #79# and of the bulk of gene-edited FA-A MEFs (**Figure 20 B** and **C**). As in some clones we could not observe TI by FISH, we have hypothesized that the basal genomic instability of these cells would have led to rearrangements and loss of the transgene integration.

In the last experiments with gene-edited FA-A MEFs, we confirmed the functionality of the integrated therapeutic cassette. By WB, we confirmed the expression of the hFANCA protein in all gene-edited clones (**Figure 21**). Furthermore, we confirmed the MMC resistance

of different gene-edited FA-A MEF clones (**Figure 22**). Additionally, when we analysed the chromosomal instability of gene-edited clones, we observed a significant basal genomic instability, an observation that could also be caused by the immortalization of FA-A MEFs (**Figure 23 A).** More significantly, we observed that all clones reverted the characteristic induction of chromosomal aberrations with MMC (**Figure 23 B**).

Taken together, our data constitute the first demonstration of a *safe harbor*-based gene-targeting strategy in FA-A MEFs. Furthermore, we have demonstrated the correction of the FA/BRCA pathway through the mTALEN-mediated targeting and the integration of a therapeutic h*FANCA* cassette in FA-A MEFs.

3. GENE-TARGETING STUDIES IN MOUSE HEMATOPOIETIC PROGENITORS

Once we proved the efficacy of our gene-targeting experiments in FA-A MEFs, we performed similar experiments in mouse HSPCs, as these cells are ideal targets to correct in FA, a monogenic disease in which the hematopoietic system is affected. Thus, the first experiments were done with the aim of demonstrating the feasibility of a gene-targeting strategy in the *Mbs85* locus of mHSPCs.

3.1. Optimization of the nucleofection required for gene targeting of mouse hematopoietic progenitors

The gene targeting of mHSPCs with nucleases is challenging. Different papers have been published about strategies of gene targeting in human HSPCs [113, 184, 216, 560]. However, in mouse, there is only one publication in which the authors have targeted primary mHSPC Lin^- $Sca\text{-}1^+$ cells [457]. Alternatively, the team of Dr. Luigi Naldini (Schiroli G et al.) is trying to target the *Il2rg* locus of mHSPCs from a mouse model of SCID-X1.

With the aim of implementing gene targeting in mouse HSPCs, Lin^- BM cells were purified by cell sorting, and then prestimulated for 48 hours to promote cell cycling of quiescent BM HSPCs. Prestimulated cells that were nucleofected with the mTALEN and the donor were then maintained in the presence of hematopoietic growth factors to boost HRR-mediated gene-editing repair, which is predominant in S/G2 phases of cell cycle. But first, we tuned up the nucleofection conditions, evaluating the effects of nucleofection upon cytotoxicity, transfection efficiency, and clonogenic potential of nucleofected Lin^- BM cells. Up to eight nucleofection programs were evaluated with different *EGFP* control DNA plasmids. As the ED-113 program was the most efficient both in terms of viability and transfection efficiency of the Lin^- BM cells (**Figure 24 A**), we decided to use this program in further experiments. Additionally, one of the features of the HSPCs is their ability to differentiate to more committed lineages. Thus, we analysed the clonogenic potential of nucleofected mHSPCs, and we observed that the survival of the hematopoietic colonies was markedly affected due to plasmid DNA nucleofection (**Figure 24 B**), contrary to what we observed when we nucleofected FA-A MEFs.

An important parameter that affects the maintenance of the multipotency of HSPCs and their differentiation potential is the presence of different growth factors in the culture media. The cocktail of cytokines that we used was chosen based on other experiments of *in vitro* culture of mHSPCs from our laboratory [386]. Different articles have been published, in which different concentrations of cytokines were used in humans. In these studies, hSCF, hTPO, and hFlt3 cytokines were included [113, 184, 216, 560]. In our particular case, although we tested different cocktails during the first experiments, we decided to use the StemSpan™ medium and 100 ng/ml of each hematopoietic factor, mSCF, hIL11, hFlt3, hTPO (**Table S4**). Furthermore, in the most recent publications working with human HSPCs [113, 560] and also in the gene-editing paper involving mHSPCs [457], the authors used 100 ng/ml of each cytokine, in accordance with our

protocol. With this cocktail, we observed that after the liquid culture of mHSPCs, we obtained a considerable number of hematopoietic colonies *in vitro* (**Figure 24 B**).

In these optimization experiments we also analysed the LSK compartment, which contains very primitive mHSPCs. Nucleofection with DNA clearly compromised this compartment. Nevertheless, we could observe that the number of transfected LSK expressing EGFP was 3.8%. This percentage was in the same range as what Genovese et al. [184] reported in the different subpopulations of targeted human primitive HSPCs. In addition, the total percentage of $EGFP^+LSK^+$ cells that survived the process (0.14%) suggested that a very reduced number of transfected progenitors were recovered after the procedure of DNA nucleofection.

Due to the observed cytotoxicity of DNA nucleofected cells with plasmid DNA, we decided to test the efficacy of nucleofecting mHSPCs with mRNA. As in the previous experiments, we first evaluated the cytotoxicity, efficiency, and clonogenic ability of the nucleofected Lin^- BM cells with different batches of mRNA that expressed EGFP. The general conclusion was that despite the use of different programs, doses of mRNA, batches of mRNA, and time points of evaluation, the expression of EGFP in these cells was very low. This low transfection efficiency inversely affected their viability in liquid culture and their clonogenic ability. We have hypothesized that mRNA nucleofection did not work because the pulse electroporation or the chemical formulations used during the process (nucleofection reagents) could degrade the mRNA before penetrating into these mHSPCs. Hence, based on these results, we decided to work with both the mTALEN and the donor as plasmid DNA using the ED-113 program for nucleofecting Lin^- BM cells.

In the next set of experiments, we analysed if the access to the target *Mbs85* locus and its cleavage by the NN-mTALEN could vary depending on the cell type, WT or FA-A Lin^- BM cells. With this purpose we performed Surveyor assays in these gene-edited cells. However, we did not observe significant differences between the different cell genotypes (**Figure 26** and **Figure 35**). This result could suggest that the NHEJ repair pathway is not significantly affected in FA-A Lin^- BM cells. This is in accordance with what Nakanishi et al. described, proposing that NHEJ repair was not reduced in *FANCA* deficient cells [381], although contrasts with other studies that indicate that the NHEJ pathway is affected in FA-A cells [63, 127, 136, 591]. Up to now, it is very controversial if the NHEJ pathway is increased or reduced in FA defective cells.

3.2. Gene-targeting analyses in WT hematopoietic progenitors with the PGK-*EGFP* reporter donor

Once we proved the efficient cleavage of the NN-mTALEN pair in the *Mbs85* locus of mouse HSPCs both from WT and FA-A mice, we decided to implement a gene-targeting approach in mHSPCs. In this set of experiments we used an *EGFP* reporter donor, constructed in the same backbone and with the same elements as that of the therapeutic h*FANCA* construct. Even under optimized conditions a very significant toxicity was observed in nucleofected mHSPCs (**Figure 27 A** and **C**). Increasing the post-nucleofection time from 2 days to 5 days prior to establishing the semisolid cultures did not improve significantly the clonogenic potential of these cells (**Figure 28**). Therefore, we decided to continue with the time point of 2 days after nucleofection for subsequent analyses.

When we analyse EGFP expression, we did not observe significant differences at 2 days post-nucleofection in samples nucleofected with the mTALEN plus the donor in comparison with nucleofected samples with only the donor (**Figure 27 A**). However, at 14 days post-nucleofection, EGFP expression was only observed in cells transfected both with the mTALEN and donor (**Figure 29**). These results suggest that the PGK-*EGFP* integration and expression was mediated by the mTALEN that cleave in the *Mbs85* locus of these cells. The MFI analysed in these samples at 14 days post-nucleofection corroborated the absence of EGFP expression in donor nucleofected samples (**Figure 30 B**). It is remarkable that cells analysed at 14 days post-nucleofection had a higher intensity of expression (MFI) as compared to values obtained at 2 days post-nucleofection. These results indicate the higher activity of the eukaryotic PGK promoter when the donor is integrated in the cell genome as compared to non-integrated copies of the donor.

The estimated TI frequencies analysed by FACS at 14 days post-nucleofection varied from 0.12% (with 0.75 µg of mTALEN and 4 µg of donor) to 0.28% (with 2.5 µg of mTALEN and 4 µg of donor). And we also calculated the TI frequencies in the individual colonies that arose from clonogenic cultures treated with the different selection methods, by conducting PCRs for the 5' and the 3' integration junctions. Since we did not observe EGFP expression in any of these colonies, we analysed most of the colonies that arose from each condition and from each experiment. Based on these analyses the mean estimated TI frequencies were in the range of 2.5% to 20.98% (considering the colonies that were positive for any of the in-out PCR analyses). TI frequencies were higher when higher doses of the mTALEN were used, probably due to their more efficient cleavage in the *Mbs85* locus. TI frequencies observed when the nucleofected cells were maintained in culture for 5 days post-nucleofection were in the same range as compared with values obtained in these cells maintained for only 2 days. In most colonies, we had difficulties amplifying the specific PCR bands corresponding for both integration junctions. This could be due to the efficiency of the PCRs, or to the limited and poor quality of the DNA obtained from single colonies. However, as has been previously mentioned, other mechanisms of repair of the DSBs could have been involved apart from HRR, giving rise to different outcomes at the different junctions. Importantly, most of the colonies that arose from samples only nucleofected with the donor did not amplify any of the specific integration bands. In the few cases in which this was observed, on-target non-directed integration events independent of mTALEN cleavage may have occurred.

TI frequencies by FACS in samples nucleofected with 4 µg of donor ranged from 0.12% to 0.28% and by the in-out PCRs, they ranged from 6.47% to 20.98%. These differences could be due to the fact that more than one copy of the cassette would be necessary to observe EGFP expression by FACS. Alternatively, the *Mbs85* locus could limit the transgene expression in these cells. In this respect, whereas in liquid culture we could observe EGFP expression in a low proportion of cells, we could not observe any EGFP-expressing hematopoietic colony. Thus, we hypothesized that $EGFP^+$ cells observed in the liquid cultures were more differentiated as compared to the CFCs responsible of the generation of the hematopoietic colonies. This would imply a restricted expression of transgenes in the *Mbs85* locus of mHSPCs.

Regarding the faithful expression of a transgene into the *Mbs85* locus, only one study has been published so far about the expression of a transgene integrated into the *Mbs85* locus

of mESCs. The study described the integration of a rAAV which allowed the expression of the *EGFP* from the rAAV [211]. Additionally, with respect to the human *AAVS1* locus, transgene expression has been demonstrated in different cell types including human ESCs (hESCs) [120, 217, 318, 437, 498] and $CD34^+$ cells [113, 184, 560]. Furthermore, the human *AAVS1* locus has been described as an open chromatin locus, which also contains a putative insulator element, facilitating the incorporation and expression of transgenes without interfering in the expression of surrounding genes [546]. However, it has been recently described that the region that the nucleases frequently target in the human *AAVS1* locus is methylated in hESCs, limiting a reliable transgene expression [400]. Whether or not this region in the *Mbs85* locus is also methylated in mHSPCs is currently unknown, although this could account for the restricted expression of EGFP in mHSPCs.

To date, it remains unclear whether the low efficiency of HDR in HSPCs is only due to their quiescent nature [360, 380] or because other processes are involved in these rare events. In our experiments with WT Lin^- BM cells, INDEL frequencies generated by NHEJ varied from 24.62% to 36.50%, and were much higher than the TI frequencies mediated by HDR (mean values in WT hematopoietic colonies ranged from 2.50% to 20.98%). This difference has been described in other studies with human HSPCs [184]. Perhaps our delivery method, based on plasmid DNA, was very detrimental for the viability of mHSPCs, which led to a lower gene-editing efficiency. Different possibilities for enhancing the gene-targeting efficiency have been described, including the inhibition of NHEJ or the enhancement of HRR. Synchronizing cells before S/G2 phases of cell cycle prior to transfection has shown to increase Cas9-induced HDR efficiency [303]. To favour the repair of the DSBs by HDR, some groups have inhibited Ligase IV (which is implicated in NHEJ pathway) with the Scr7 inhibitor or the coexpression of adenovirus 4 E1B55K and E4orf6 proteins [105, 334]. Other approaches, such as the use of nickases, have been described as reducing NHEJ events, because they are legitimate substrates of SSA repair pathway, and promote higher levels of error-free HDR compared with NHEJ-mediated insertions/deletions [403]. The inclusion of other factors, such as LEDGF (p75) that favours the recruitment of CtIP to DNA DSBs could also be a potential strategy to promote HDR in gene-targeting approaches [109]. Finally, the use of the inhibitor antirecombinase PARI to increase the efficiency of HRR in FA fibroblasts has already been proved with moderate efficacies [364].

In this respect, within the last year, several molecules, such as L755507 and BrefeldinA were proposed to enhance HDR in mESCs. Furthermore, L755507 consistently increased HDR in other cell types, such as K562, HeLa, HUVEC, CRL-2097 fibroblast and hESCs-derived neural stem cells. In this thesis, we studied the effect of these molecules on promoting TI in WT Lin^- BM cells. Since only one experiment was performed with these molecules, we can not reach a definitive conclusion about these treatments. However, TI was not improved by the treatments with these molecules. This result is in accordance with a recent study in which it was reported that no significant increases in HDR efficiency were induced by L755507 alone, but it was effective when combined with RS-1 [429, 502].

Our TI frequencies calculated in WT hematopoietic colonies by in-out PCRs are quite a bit lower than the latest gene-targeting approaches in $CD34^+$ using ZFNs. The authors obtained TI frequencies that ranged from 10% to 26% [184 113, 216, 560]. Regarding the use of TALEN to target

HSPCs, one study related to the generation of leukemia fusion oncogenes showed an average of 9% knock-in efficiency in $CD34^+$ cells from healthy donors [51]. Furthermore, there is a publication in which the authors implemented a gene-editing strategy in mouse Lin^- $Sca\text{-}1^+$ cells derived from an *Artemis* SCID mouse model, using both the *I-Sce1* meganuclease and the donor delivered by SIN-IDLV. The authors gave an estimation frequency of targeted integration based on T-cell differentiation of 2.04% and 1.95% in two independent experiments [457]. And lastly, Schiroli et al. obtained 6% transgene integration by HDR delivering the donor template by IDLVs and the ZFNs as mRNA in mouse *Il2rg*$^{-/-}$ HSPCs. Upon transplant, the corrected cells showed reconstitution of only the lymphoid lineage compartment of lethally irradiated mice (unpublished results, oral communication OR019, ESGCT 2015). These publications, combined with our data, indicate that apart from the cytotoxicity of the procedure mediated by plasmid DNA nucleofection, the mHSPC population constitute a difficult target in which to implement a targeted integration approach using TALEN, perhaps due to the very low activity of the HR pathway in these cells.

Finally, in an attempt to investigate the efficiency in the gene editing of the very primitive mHSPCs, we performed some experiments conducting transplants using C57BL/6 WT mice as donors ($CD45.2^+$) and P3B WT ($CD45.1^+$) mice as recipients. In the transplant analyses performed with nucleofected WT Lin^- BM cells, we observed a marked compromise of the repopulating potential of DNA nucleofected cells in the irradiated recipients (**Figure 34** and **Table S5**). This defect in the repopulating potential was neither improved by the maintenance of the cells in culture for 5 days after nucleofection nor by the use of molecules that could enhance the frequency of HDR events. Thus, the defect in the *in vivo* repopulating properties of DNA nucleofected WT Lin^- BM cells according to our conditions, did not allow us to demonstrate whether this very primitive population had been successfully edited.

3.3. Gene-targeting analyses in FA-A hematopoietic progenitors with the therapeutic PGK-h*FANCA* donor

FA has been proposed as an appropriate disease to implement gene-targeting approaches due to the proliferative advantage of the corrected HSPCs over the non-corrected ones, as it has been proposed that this effect could result in a progressive restoration of the hematopoietic lineage in FA patients. Moreover, this advantage has been already observed in mosaic FA patients. Nevertheless, the main limitation in a gene therapy approach in FA patients resides in the limited pool of HSCs that FA patients contain, which thus restrict the number of corrected cells that can be infused into the them.

In our experiments with FA-A mHSPCs, we did not observe differences in the viability, transfection efficiency and clonogenic potential of nucleofected FA-A mHSPCs (**Figure 36 A** and **B)** in comparison with experiments performed in WT mHSPCs (**Figure 27 A and C**). However, there was a trend indicating that nucleofected WT Lin^- BM cells presented higher viabilities than their FA-A counterparts. The higher viabilities observed in nucleofected WT mHSPCs could suggest that smaller donors, such as the *EGFP* reporter donor, were less toxic for the viability of WT Lin^- BM cells. Nevertheless, this could also be explained by the differences between the two C57BL/6 WT and FVB FA-A cell genotypes.

Clonogenic cultures selected either with puromycin (**Figure 36 C**) or with MMC (**Figure 37**) demonstrated an increase in the survival of the colonies that arose from FA-A mHSPCs

nucleofected with both the mTALEN plus donor. This was more evident in the condition where 4 µg of donor with the mTALEN were used in contrast with the use of 4 µg of donor alone. These results indicate that random integration events (donor alone) were less frequent compared to on-target integration events (mTALEN plus donor). Furthermore, the observation that gene-edited FA-A cells (nucleofected with the mTALEN with the donor) survived cytotoxic concentrations of MMC is consistent with data obtained in FA-A MEFs, and indicates the therapeutic effect of the proposed gene therapy approach.

In subsequent experiments, we investigated if the integration of the therapeutic PGK-h*FANCA* cassette had occurred and at what efficiencies into the *Mbs85* locus of FA-A Lin$^-$ BM cells. We observed different integration frequencies depending on the selection method that we used (**Figure 38**). Higher percentages of gene editing where obtained with the dose of 4 µg of donor (**Figure 41**). Moreover, as was observed in hematopoietic colonies derived from WT mHSPCs, only a low proportion of colonies derived from gene-edited FA-A mHSPCs amplified both bands corresponding to the 5' or the 3' integration junctions. This could be due to any of the previous explanations that we have already mentioned in the section of the discussion related to targeted integration experiments in WT hematopoietic progenitors.

The comparison of the TI frequencies observed in hematopoietic colonies from WT and FA-A Lin$^-$ BM revealed TI frequencies lower in FA-A cells (**Figure 31** and **Figure 41**). If we would not have taken into account the viabilities after selection, the TI percentage both in FA-A and in WT HSPCs would be in the same range. Due to this, we could hypothesize that the selection methods used in the experiments performed in FA-A Lin$^-$ BM cells could not be selecting but leading the cells to cell death, thereby, reducing the overall TI percentages. Whether or not these differences are due to the permissiveness of the cell strain, due to an impaired HRR in FA-A cells, or due to the differences in the donor's size can not be known from our results. However, in recent work developed by our laboratory (B. Díez-Cabezas et al.) the authors observed that TI frequencies measured by FACS were normally higher in LCLs from healthy donors than in LCLs from FA-A patients. These data would suggest that although FANCA is not essential in the repair of DSBs by HDR, it could somehow be implicated in the repair efficiency by HR.

Considering that our final goal was to restore the hematopoiesis of FA-A mice, we tested the repopulating ability of gene-edited FA-A Lin$^-$ BM cells in myeloablated female FA-A mice. All transplanted mice died, probably due to graft failure by reason of the few numbers of the nucleofected gene-edited cells that were transplanted. Nevertheless, to further investigate this possibility we analysed what the minimum number of untransfected cells that could engraft FA-A mice was. We observed that with almost all the different doses of cells tested (from $3x10^5$ to $6x10^5$, which were in the range of cells that we obtained in the nucleofection experiments) the engrafted mice died or presented very low levels of engraftment. The transplant of FA-A Lin$^-$ BM cells in FA-A recipients manifested a considerable problem of hematopoietic graft failure (**Figure 42**). This engraftment defect has been already described in other mouse models such as *Fancc*$^{-/-}$ [205] or *Fancd1* (*Brca2*$^{\Delta 27/\Delta 27}$) [386]. These results suggest that without correcting the FA-A Lin$^-$ BM cells, the probability of graft failure is very high. Thus, higher doses of gene-edited cells in comparison to what we have been used are needed to repopulate the BM of FA-A recipient mice. Moreover, as we could not prove the engraftment

of gene-edited FA-A Lin^- BM cells, we can not confirm that these cells were the most primitive mHSPCs, which are the ones with long-term engraftment capabilities. This could be because, as has been described, the repair of DSBs by HDR is more active in more mature cells and less in more primitive mHSPCs [216, 360].

Altogether, these results demonstrate that gene targeting in mHSPCs is challenging. Better delivery platforms could probably improve the efficiency of the strategy. However, the fact that there is only one published work in this field with mHSPCs suggests that other obstacles impede the development of gene-editing strategies in mouse hematopoietic progenitor cells.

X. CONCLUSIONS

1. We have demonstrated the activity of new TALEN designed to cleave the mouse *Mbs85* locus, an ortholog of the human *AAVS1 safe harbor* locus.

2. We have generated two donor cassettes either containing the *EGFP* reporter gene or the therapeutic h*FANCA* gene, both under the regulation of the PGK promoter and flanked by two arms homologous to the *Mbs85* locus sequences.

3. We have developed a therapeutic gene-editing strategy in FA-A MEFs that reverted the characteristic MMC hypersensitivity of FA-A cells through the expression of the hFANCA protein.

4. We have demonstrated the feasibility of implementing a gene-targeting approach into the *Mbs85* locus of WT and FA-A mouse hematopoietic progenitor cells.

5. Using a therapeutic gene-targeting strategy with the h*FANCA* donor, we have demonstrated the reversion of the characteristic phenotype of FA-A hematopoietic progenitor cells.

6. We have observed that the therapeutic potential of the gene-editing strategy is limited because of the compromised repopulating ability of WT and FA-A HSCs exposed to DNA nucleofection.

Taken together, these results demonstrate for the first time the feasibility of conducting a therapeutic gene-targeting strategy in the *Mbs85 safe harbour* locus of embryonic fibroblasts and hematopoietic progenitors from a mouse model of FA-A. Further studies should be performed aiming at reducing the toxicity associated to the delivery of the nucleases and donor sequences in HSPCs, in order to increase their repopulating properties.

XI. CONCLUSIONES

1. Hemos demostrado la actividad de nuevas TALEN diseñadas para cortar el locus *Mbs85* de ratón, ortólogo al *locus seguro AAVS1* humano.

2. Hemos generado dos construcciones donadoras que contienen el gen marcador *EGFP* o el gen terapéutico h*FANCA*, ambos regulados por el promotor PGK y flanqueados por dos brazos homólogos a secuencias del locus *Mbs85*.

3. Hemos desarrollado una estrategia terapéutica de edición génica en fibroblastos embrionarios AF-A que revirtió la hipersensibilidad a MMC característica de células de AF-A mediante la expresión de la proteína hFANCA.

4. Hemos demostrado la posibilidad de llevar a cabo una aproximación de terapia génica dirigida en el locus *Mbs85* de progenitores hematopoyéticos de ratones sanos (WT) y de ratones AF-A.

5. Usando una estrategia de terapia génica dirigida con un donador h*FANCA,* hemos demostrado la reversión del fenotipo característico de progenitores hematopoyéticos AF-A.

6. Hemos observado que el potencial terapéutico de esta estrategia de edición génica es limitado al comprometerse la capacidad de repoblación de las CMHs de ratones sanos (WT) y de AF-A expuestas a la nucleofección de ADN.

En conjunto, estos resultados demuestran por primera vez la posibilidad de llevar a cabo una estrategia terapéutica de terapia génica dirigida en el *locus seguro* de ratón *Mbs*85 de fibroblastos embrionarios y progenitores hematopoyéticos de ratones modelo de AF-A. Estudios adicionales deberán llevarse a cabo con el objetivo de reducir la toxicidad asociada a la introducción de las nucleasas y las secuencias donadoras en CMPHs, con el propósito de potenciar sus propiedades de repoblación.

XII. BIBLIOGRAPHY

1 Adamo A, Collis SJ, Adelman CA, Silva N, Horejsi Z, Ward JD *et al*. Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. Molecular cell 2010; 39: 25-35.

2 Agarwal S, Tafel AA, Kanaar R. DNA double-strand break repair and chromosome translocations. DNA Repair (Amst) 2006; 5: 1075-1081.

3 Agoulnik AI, Lu B, Zhu Q, Truong C, Ty MT, Arango N *et al*. A novel gene, Pog, is necessary for primordial germ cell proliferation in the mouse and underlies the germ cell deficient mutation, gcd. Hum Mol Genet 2002; 11: 3047-3053.

4 Aiuti A, Slavin S, Aker M, Ficara F, Deola S, Mortellaro A *et al*. Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. Science 2002; 296: 2410-2413.

5 Aiuti A, Cattaneo F, Galimberti S, Benninghoff U, Cassani B, Callegaro L *et al*. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med 2009; 360: 447-458.

6 Aiuti A, Biasco L, Scaramuzza S, Ferrua F, Cicalese MP, Baricordi C *et al*. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science 2013; 341: 1233151.

7 Ali AM, Pradhan A, Singh TR, Du C, Li J, Wahengbam K *et al*. FAAP20: a novel ubiquitin-binding FA nuclear core-complex protein required for functional integrity of the FA-BRCA DNA repair pathway. Blood 2012; 119: 3285-3294.

8 Alter BP, Kupfer G. Fanconi Anemia 1993.

9 Aly A, Ganesan S. BRCA1, PARP, and 53BP1: conditional synthetic lethality and synthetic viability. J Mol Cell Biol 2011; 3: 66-74.

10 Amendola M, Venneri MA, Biffi A, Vigna E, Naldini L. Coordinate dual-gene transgenesis by lentiviral vectors carrying synthetic bidirectional promoters. Nat Biotechnol 2005; 23: 108-116.

11 Andreassen PR, D'Andrea AD, Taniguchi T. ATR couples FANCD2 monoubiquitination to the DNA-damage response. Genes Dev 2004; 18: 1958-1963.

12 Annoni A, Goudy K, Akbarpour M, Naldini L, Roncarolo MG. Immune responses in liver-directed lentiviral gene therapy. Transl Res 2013; 161: 230-240.

13 Antonio Casado J, Callen E, Jacome A, Rio P, Castella M, Lobitz S *et al*. A comprehensive strategy for the subtyping of patients with Fanconi anaemia: conclusions from the Spanish Fanconi Anemia Research Network. J Med Genet 2007; 44: 241-249.

14 Aronin N, DiFiglia M. Huntingtin-lowering strategies in Huntington's disease: antisense oligonucleotides, small RNAs, and gene editing. Mov Disord 2014; 29: 1455-1461.

15 Asokan A, Schaffer DV, Samulski RJ. The AAV vector toolkit: poised at the clinical crossroads. Mol Ther 2012; 20: 699-708.

16 Auerbach AD, Rogatko A, Schroeder-Kurth TM. International Fanconi Anemia Registry: relation of clinical symptoms to diepoxybutane sensitivity. Blood 1989; 73: 391-396.

17 Auerbach AD, Liu Q, Ghosh R, Pollack MS, Douglas GW, Broxmeyer HE. Prenatal identification of potential donors for umbilical cord blood transplantation for Fanconi anemia. Transfusion 1990; 30: 682-687.

18 Auerbach AD. Fanconi anemia and its diagnosis. Mutat Res 2009; 668: 4-10.

19 Bae KH, Kwon YD, Shin HC, Hwang MS, Ryu EH, Park KS *et al*. Human zinc fingers as building blocks in the construction of artificial transcription factors. Nat Biotechnol 2003; 21: 275-280.

20 Baker SM, Plug AW, Prolla TA, Bronner CE, Harris AC, Yao X *et al*. Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over. Nat Genet 1996; 13: 336-342.

21 Bakker ST, van de Vrugt HJ, Rooimans MA, Oostra AB, Steltenpool J, Delzenne-Goette E *et al*. Fancm-deficient mice reveal unique features of Fanconi anemia complementation group M. Hum Mol Genet 2009; 18: 3484-3495.

22 Bakker ST, van de Vrugt HJ, Visser JA, Delzenne-Goette E, van der Wal A, Berns MA *et al*. Fancf-deficient mice are prone to develop ovarian tumours. J Pathol 2012; 226: 28-39.

23 Bakker ST, de Winter JP, te Riele H. Learning from a paradox: recent insights into Fanconi anaemia through studying mouse models. Dis Model Mech 2013; 6: 40-47.

24 Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S *et al*. CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007; 315: 1709-1712.

25 Barrett DM, Grupp SA, June CH. Chimeric Antigen Receptor- and TCR-Modified T Cells Enter Main Street and Wall Street. J Immunol 2015; 195: 755-761.

26 Baum C, von Kalle C, Staal FJ, Li Z, Fehse B, Schmidt M *et al*. Chance or necessity? Insertional mutagenesis in gene therapy and its consequences. Mol Ther 2004; 9: 5-13.

27 Becker PS, Taylor JA, Trobridge GD, Zhao X, Beard BC, Chien S *et al*. Preclinical correction of human Fanconi anemia complementation group A bone marrow cells using a safety-modified lentiviral vector. Gene Ther 2010; 17: 1244-1252.

28 Bessis N, GarciaCozar FJ, Boissier MC. Immune responses to gene therapy vectors: influence on vector function and effector mechanisms. Gene Ther 2004; 11 Suppl 1: S10-17.

29 Bianchi M, Hakkim A, Brinkmann V, Siler U, Seger RA, Zychlinsky A *et al*. Restoration of NET formation by gene therapy in CGD controls aspergillosis. Blood 2009; 114: 2619-2622.

30 Bianchi M, Niemiec MJ, Siler U, Urban CF, Reichenbach J. Restoration of anti-Aspergillus defense by neutrophil extracellular traps in human chronic granulomatous disease after gene therapy is calprotectin-dependent. The Journal of allergy and clinical immunology 2011; 127: 1243-1252 e1247.

31 Bielorai B, Hughes MR, Auerbach AD, Nagler A, Loewenthal R, Rechavi G *et al*. Successful umbilical cord blood transplantation for Fanconi anemia using preimplantation genetic diagnosis for HLA-matched donor. Am J Hematol 2004; 77: 397-399.

32 Biffi A, Montini E, Lorioli L, Cesani M, Fumagalli F, Plati T *et al*. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. Science 2013; 341: 1233158.

33 Blackford AN, Schwab RA, Nieminuszczy J, Deans AJ, West SC, Niedzwiedz W. The DNA translocase activity of FANCM protects stalled replication forks. Hum Mol Genet 2012; 21: 2005-2016.

34 Blaese RM, Culver KW, Miller AD, Carter CS, Fleisher T, Clerici M *et al*. T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. Science 1995; 270: 475-480.

35 Bluteau D, Masliah-Planchon J, Clairmont C, Rousseau A, Ceccaldi R, Dubois d'Enghien C *et al*. Biallelic inactivation of REV7 is associated with Fanconi anemia. J Clin Invest 2016.

36 Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annual review of phytopathology 2010; 48: 419-436.

37 Bogdanove AJ, Schornack S, Lahaye T. TAL effectors: finding plant genes for disease and defense. Current opinion in plant biology 2010; 13: 394-401.

38 Bogliolo M, Schuster B, Stoepker C, Derkunt B, Su Y, Raams A *et al*. Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. American journal of human genetics 2013; 92: 800-806.

39 Bogliolo M, Surralles J. Fanconi anemia: a model disease for studies on human genetics and advanced therapeutics. Curr Opin Genet Dev 2015; 33: 32-40.

40 Bonfim CM, de Medeiros CR, Bitencourt MA, Zanis-Neto J, Funke VA, Setubal DC *et al*. HLA-matched related donor hematopoietic cell transplantation in 43 patients with Fanconi anemia conditioned with 60 mg/kg of cyclophosphamide. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation 2007; 13: 1455-1460.

41 Booth C, Gaspar HB, Thrasher AJ. Treating Immunodeficiency through HSC Gene Therapy. Trends Mol Med 2016; 22: 317-327.

42 Bosch PC, Bogliolo M, Surralles J. Activation of the Fanconi anemia/BRCA pathway at low doses of ionization radiation. Mutat Res Genet Toxicol Environ Mutagen 2015; 793: 9-13.

43 Bothmer A, Robbiani DF, Feldhahn N, Gazumyan A, Nussenzweig A, Nussenzweig MC. 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. The Journal of experimental medicine 2010; 207: 855-865.

44 Boussif O, Lezoualc'h F, Zanta MA, Mergny MD, Scherman D, Demeneix B *et al*. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. Proc Natl Acad Sci U S A 1995; 92: 7297-7301.

45 Bouwman P, Aly A, Escandell JM, Pieterse M, Bartkova J, van der Gulden H *et al*. 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. Nat Struct Mol Biol 2010; 17: 688-695.

46 Bouwman P, Drost R, Klijn C, Pieterse M, van der Gulden H, Song JY *et al*. Loss of p53 partially rescues embryonic development of Palb2 knockout mice but does not foster haploinsufficiency of Palb2 in tumour suppression. J Pathol 2011; 224: 10-21.

47 Boztug K, Schmidt M, Schwarzer A, Banerjee PP, Diez IA, Dewey RA *et al*. Stem-cell gene therapy for the Wiskott-Aldrich syndrome. N Engl J Med 2010; 363: 1918-1927.

48 Braun CJ, Boztug K, Paruzynski A, Witzel M, Schwarzer A, Rothe M *et al*. Gene therapy for Wiskott-Aldrich syndrome--long-term efficacy and genotoxicity. Science translational medicine 2014; 6: 227ra233.

49 Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P *et al*. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res 2012; 40: e117.

50 Brown BD, Gentner B, Cantore A, Colleoni S, Amendola M, Zingale A *et al*. Endogenous microRNA can be broadly exploited to regulate transgene expression according to tissue, lineage and differentiation state. Nat Biotechnol 2007; 25: 1457-1467.

51 Buechele C, Breese EH, Schneidawind D, Lin CH, Jeong J, Duque-Afonso J *et al*. MLL leukemia induction by genome editing of human CD34+ hematopoietic cells. Blood 2016; 126: 1683-1694.

52 Bunting SF, Callen E, Wong N, Chen HT, Polato F, Gunn A *et al*. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. Cell 2010; 141: 243-254.

53 Bunting SF, Nussenzweig A. Dangerous liaisons: Fanconi anemia and toxic nonhomologous end joining in DNA crosslink repair. Molecular cell 2010; 39: 164-166.

54 Cai Y, Bak RO, Mikkelsen JG. Targeted genome editing by lentiviral protein transduction of zinc-finger and TAL-effector nucleases. Elife 2014; 3: e01911.

55 Callen E, Samper E, Ramirez MJ, Creus A, Marcos R, Ortega JJ *et al*. Breaks at telomeres and TRF2-independent end fusions in Fanconi anemia. Hum Mol Genet 2002; 11: 439-444.

56 Callen E, Casado JA, Tischkowitz MD, Bueren JA, Creus A, Marcos R *et al*. A common founder mutation in FANCA underlies the world's highest prevalence of Fanconi anemia in Gypsy families from Spain. Blood 2005; 105: 1946-1949.

57 Candotti F, Shaw KL, Muul L, Carbonaro D, Sokolic R, Choi C *et al*. Gene therapy for adenosine deaminase-deficient severe combined immune deficiency: clinical comparison of retroviral vectors and treatment plans. Blood 2012; 120: 3635-3646.

58 Canver MC, Smith EC, Sher F, Pinello L, Sanjana NE, Shalem O *et al*. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature 2015; 527: 192-197.

59 Capecchi MR. Altering the genome by homologous recombination. Science 1989; 244: 1288-1292.

60 Carreau M. Not-so-novel phenotypes in the Fanconi anemia group D2 mouse model. Blood 2004; 103: 2430.

61 Carroll D. Staying on target with CRISPR-Cas. Nat Biotechnol 2013; 31: 807-809.

62 Cartier N, Hacein-Bey-Abina S, Bartholomae CC, Veres G, Schmidt M, Kutschera I *et al*. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. Science 2009; 326: 818-823.

63 Casado JA, Nunez MI, Segovia JC, Ruiz de Almodovar JM, Bueren JA. Non-homologous end-joining defect in fanconi anemia hematopoietic cells exposed to ionizing radiation. Radiation research 2005; 164: 635-641.

64 Casado JA, Callen E, Jacome A, Rio P, Castella M, Lobitz S *et al*. A comprehensive strategy for the subtyping of Fanconi Anemia patients: conclusions from the Spanish Fanconi Anemia research network. J Med Genet 2007; 44: 241-249.

65 Caserta TM, Smith AN, Gultice AD, Reedy MA, Brown TL. Q-VD-OPh, a broad spectrum caspase inhibitor with potent antiapoptotic properties. Apoptosis 2003; 8: 345-352.

66 Castiello MC, Scaramuzza S, Pala F, Ferrua F, Uva P, Brigida I *et al*. B-cell reconstitution after lentiviral vector-mediated gene therapy in patients with Wiskott-Aldrich syndrome. The Journal of allergy and clinical immunology 2015; 136: 692-702 e692.

67 Cavazzana-Calvo M, Payen E, Negre O, Wang G, Hehir K, Fusil F *et al*. Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. Nature 2010; 467: 318-322.

68 Ceccaldi R, Parmar K, Mouly E, Delord M, Kim JM, Regairaz M *et al*. Bone marrow failure in Fanconi anemia is triggered by an exacerbated p53/p21 DNA damage response that impairs hematopoietic stem and progenitor cells. Cell Stem Cell 2012; 11: 36-49.

69 Ceccaldi R, Rondinelli B, D'Andrea AD. Repair Pathway Choices and Consequences at the Double-Strand Break. Trends Cell Biol 2016; 26: 52-64.

70 Cesana D, Sgualdino J, Rudilosso L, Merella S, Naldini L, Montini E. Whole transcriptome characterization of aberrant splicing events induced by lentiviral vector integrations. J Clin Invest 2012; 122: 1667-1676.

71 Cicalese MP, Aiuti A. Clinical applications of gene therapy for primary immunodeficiencies. Hum Gene Ther 2015; 26: 210-219.

72 Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. Molecular cell 2010; 40: 179-204.

73 Clauson C, Scharer OD, Niedernhofer L. Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. Cold Spring Harb Perspect Biol 2013; 5: a012732.

74 Cleveland DW, Fischer SG, Kirschner MW, Laemmli UK. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J Biol Chem 1977; 252: 1102-1106.

75 Clouaire T, Legube G. DNA double strand break repair pathway choice: a chromatin based decision? Nucleus 2015; 6: 107-113.

76 Collis SJ, Ciccia A, Deans AJ, Horejsi Z, Martin JS, Maslen SL *et al*. FANCM and FAAP24 function in ATR-mediated checkpoint signaling independently of the Fanconi anemia core complex. Molecular cell 2008; 32: 313-324.

77 Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun 2012; 3: 968.

78 Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N *et al*. Multiplex genome engineering using CRISPR/Cas systems. Science 2013; 339: 819-823.

79 Costa RM, Chigancas V, Galhardo Rda S, Carvalho H, Menck CF. The eukaryotic nucleotide excision repair pathway. Biochimie 2003; 85: 1083-1099.

80 Costantino L, Sotiriou SK, Rantala JK, Magin S, Mladenov E, Helleday T *et al*. Break-induced replication repair of damaged forks induces genomic duplications in human cells. Science 2014; 343: 88-91.

81 Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet 2001; 2: 292-301.

82 Cross D, Burmester JK. Gene therapy for cancer treatment: past, present and future. Clin Med Res 2006; 4: 218-227.

83 Crossan GP, van der Weyden L, Rosado IV, Langevin F, Gaillard PH, McIntyre RE *et al*. Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. Nat Genet 2011; 43: 147-152.

84 Chan KL, Palmai-Pallag T, Ying S, Hickson ID. Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. Nat Cell Biol 2009; 11: 753-760.

85 Chandrasegaran S, Carroll D. Origins of Programmable Nucleases for Genome Engineering. Journal of molecular biology 2016; 428: 963-989.

86 Chandrasekharappa SC, Lach FP, Kimble DC, Kamat A, Teer JK, Donovan FX *et al*. Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia. Blood 2013; 121: e138-148.

87 Chang DJ, Cimprich KA. DNA damage tolerance: when it's OK to make mistakes. Nat Chem Biol 2009; 5: 82-90.

88 Chang L, Yuan W, Zeng H, Zhou Q, Wei W, Zhou J *et al*. Whole exome sequencing reveals concomitant mutations of multiple FA genes in individual Fanconi anemia patients. BMC Med Genomics 2014; 7: 24.

89 Chapman JR, Taylor MR, Boulton SJ. Playing the end game: DNA double-strand break repair pathway choice. Mol Cell 2012; 47: 497-510.

90 Charrier S, Ferrand M, Zerbato M, Precigout G, Viornery A, Bucher-Laurent S *et al*. Quantification of lentiviral vector copy numbers in individual hematopoietic colony-forming cells shows vector dose-dependent effects on the frequency and level of transduction. Gene Ther 2011; 18: 479-487.

91 Chen CH, Ferreira JC, Gross ER, Mochly-Rosen D. Targeting aldehyde dehydrogenase 2: new therapeutic opportunities. Physiol Rev 2014; 94: 1-34.

92 Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J *et al*. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. Nat Methods 2011; 8: 753-755.

93 Chen M, Tomkins DJ, Auerbach W, McKerlie C, Youssoufian H, Liu L *et al*. Inactivation of Fac in mice produces inducible chromosomal instability and reduced fertility reminiscent of Fanconi anaemia. Nat Genet 1996; 12: 448-451.

94 Chen X, Goncalves MA. Engineered Viruses as Genome Editing Devices. Mol Ther 2015; 24: 447-457.

95 Chen X, Bosques L, Sung P, Kupfer GM. A novel role for non-ubiquitinated FANCD2 in response to hydroxyurea-induced DNA damage. Oncogene 2016; 35: 22-34.

96 Chen ZY, He CY, Ehrhardt A, Kay MA. Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. Mol Ther 2003; 8: 495-500.

97 Cheng NC, van de Vrugt HJ, van der Valk MA, Oostra AB, Krimpenfort P, de Vries Y *et al*. Mice with a targeted disruption of the Fanconi anemia homolog Fanca. Hum Mol Genet 2000; 9: 1805-1811.

98 Chevalier BS, Stoddard BL. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. Nucleic Acids Res 2001; 29: 3757-3774.

99 Chinen J, Davis J, De Ravin SS, Hay BN, Hsu AP, Linton GF *et al*. Gene therapy improves immune function in preadolescents with X-linked severe combined immunodeficiency. Blood 2007; 110: 67-73.

100 Cho SW, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 2013; 31: 230-232.

101 Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S *et al*. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome research 2014; 24: 132-141.

102 Choulika A, Perrin A, Dujon B, Nicolas JF. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. Mol Cell Biol 1995; 15: 1968-1973.

103 Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A *et al*. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 2010; 186: 757-761.

104 Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y *et al*. Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. PLoS One 2012; 7: e45383.

105 Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K *et al*. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol 2015; 33: 543-548.

106 D'Andrea AD, Grompe M. The Fanconi anaemia/BRCA pathway. Nat Rev Cancer 2003; 3: 23-34.

107 Dang JM, Leong KW. Natural polymers for gene delivery and tissue engineering. Adv Drug Deliv Rev 2006; 58: 487-499.

108 Das SK, Menezes ME, Bhatia S, Wang XY, Emdad L, Sarkar D *et al*. Gene Therapies for Cancer: Strategies, Challenges and Successes. Journal of cellular physiology 2016; 230: 259-271.

109 Daugaard M, Baude A, Fugger K, Povlsen LK, Beck H, Sorensen CS *et al*. LEDGF (p75) promotes DNA-end resection and homologous recombination. Nat Struct Mol Biol 2012; 19: 803-810.

110 Davies AA, Huttner D, Daigaku Y, Chen S, Ulrich HD. Activation of ubiquitin-dependent DNA damage bypass is mediated by replication protein a. Molecular cell 2008; 29: 625-636.

111 Davis AJ, Chen DJ. DNA double strand break repair via non-homologous end-joining. Transl Cancer Res 2013; 2: 130-143.

112 De Palma M, Montini E, Santoni de Sio FR, Benedicenti F, Gentile A, Medico E *et al*. Promoter trapping reveals significant differences in integration site selection between MLV and HIV vectors in primary hematopoietic cells. Blood 2005; 105: 2307-2315.

113 De Ravin SS, Reik A, Liu PQ, Li L, Wu X, Su L *et al*. Targeted gene addition in human CD34(+) hematopoietic cells for correction of X-linked chronic granulomatous disease. Nat Biotechnol 2016; 34: 424-429.

114 de Winter JP, Waisfisz Q, Rooimans MA, van Berkel CG, Bosnoyan-Collins L, Alon N *et al*. The Fanconi anaemia group G gene FANCG is identical with XRCC9. Nat Genet 1998; 20: 281-283.

115 de Winter JP, Leveille F, van Berkel CG, Rooimans MA, van Der Weel L, Steltenpool J *et al*. Isolation of a cDNA representing the Fanconi anemia complementation group E gene. American journal of human genetics 2000; 67: 1306-1308.

116 de Winter JP, Rooimans MA, van Der Weel L, van Berkel CG, Alon N, Bosnoyan-Collins L *et al*. The Fanconi anaemia gene FANCF encodes a novel protein with homology to ROM. Nat Genet 2000; 24: 15-16.

117 Deans AJ, West SC. FANCM connects the genome instability disorders Bloom's Syndrome and Fanconi Anemia. Molecular cell 2009; 36: 943-953.

118 Deans AJ, West SC. DNA interstrand crosslink repair and cancer. Nat Rev Cancer 2011; 11: 467-480.

119 Decottignies A. Alternative end-joining mechanisms: a historical perspective. Front Genet 4: 48.

120 DeKelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijsing SH *et al*. Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. Genome research 2010; 20: 1133-1142.

121 Delacote F, Perez C, Guyot V, Duhamel M, Rochon C, Ollivier N *et al*. High frequency targeted mutagenesis using engineered endonucleases and DNA-end processing enzymes. PLoS One 2013; 8: e53217.

122 Deriano L, Roth DB. Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. Annu Rev Genet 2013; 47: 433-455.

123 Di Matteo M, Samara-Kuko E, Ward NJ, Waddington SN, McVey JH, Chuah MK *et al*. Hyperactive piggyBac transposons for sustained and robust liver-targeted gene therapy. Mol Ther 2014; 22: 1614-1624.

124 Ding Y, Jiang Z, Saha K, Kim CS, Kim ST, Landis RF *et al*. Gold nanoparticles for nucleic acid delivery. Mol Ther 2014; 22: 1075-1083.

125 DiPersio JF, Stadtmauer EA, Nademanee A, Micallef IN, Stiff PJ, Kaufman JL *et al*. Plerixafor and G-CSF versus placebo and G-CSF to mobilize hematopoietic stem cells for autologous stem cell transplantation in patients with multiple myeloma. Blood 2009; 113: 5720-5726.

126 Dokal I. Fanconi's anaemia and related bone marrow failure syndromes. Br Med Bull 2006; 77-78: 37-53.

127 Donahue SL, Campbell C. A DNA double strand break repair defect in Fanconi anemia fibroblasts. J Biol Chem 2002; 277: 46243-46247.

128 Donehower LA, Harvey M, Slagle BL, McArthur MJ, Montgomery CA, Jr., Butel JS *et al*. Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature 1992; 356: 215-221.

129 Dorsman JC, Levitus M, Rockx D, Rooimans MA, Oostra AB, Haitjema A *et al*. Identification of the Fanconi anemia complementation group I gene, FANCI. Cellular oncology : the official journal of the International Society for Cellular Oncology 2007; 29: 211-218.

130 Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF *et al*. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods 2011; 8: 74-79.

131 Dreyer AK, Hoffmann D, Lachmann N, Ackermann M, Steinemann D, Timm B *et al*. TALEN-mediated functional correction of X-linked chronic granulomatous disease in patient-derived induced pluripotent stem cells. Biomaterials 2015; 69: 191-200.

132 Du W, Adam Z, Rani R, Zhang X, Pang Q. Oxidative stress in Fanconi anemia hematopoiesis and disease progression. Antioxid Redox Signal 2008; 10: 1909-1921.

133 Du W, Rani R, Sipple J, Schick J, Myers KC, Mehta P *et al*. The FA pathway counteracts oxidative stress through selective protection of antioxidant defense gene promoters. Blood 2012; 119: 4142-4151.

134 Du W, Erden O, Pang Q. TNF-alpha signaling in Fanconi anemia. Blood Cells Mol Dis 2014; 52: 2-11.

135 Du W, Amarachintha S, Erden O, Wilson A, Meetei AR, Andreassen PR *et al*. Fancb deficiency impairs hematopoietic stem cell function. Sci Rep 2015; 5: 18127.

136 Du W, Amarachintha S, Wilson AF, Pang Q. Hyper-active non-homologous end joining selects for synthetic lethality resistant and pathological Fanconi anemia hematopoietic stem and progenitor cells. Sci Rep 2016; 6: 22167.

137 Dufour C, Corcione A, Svahn J, Haupt R, Poggi V, Beka'ssy AN *et al*. TNF-alpha and IFN-gamma are overexpressed in the bone marrow of Fanconi anemia patients and TNF-alpha suppresses erythropoiesis in vitro. Blood 2003; 102: 2053-2059.

138 Dufour C, Svahn J. Fanconi anaemia: new strategies. Bone Marrow Transplant 2008; 41 Suppl 2: S90-95.

139 Dumitriu B, Young NS. Damage control and its costs: BM failure in Fanconi anemia stems from overactive p53/p21. Cell Stem Cell 2012; 11: 7-8.

140 Dunn J, Potter M, Rees A, Runger TM. Activation of the Fanconi anemia/BRCA pathway and recombination repair in the cellular response to solar ultraviolet light. Cancer Res 2006; 66: 11140-11147.

141 Dupre L, Marangoni F, Scaramuzza S, Trifari S, Hernandez RJ, Aiuti A *et al*. Efficacy of gene therapy for Wiskott-Aldrich syndrome using a WAS promoter/cDNA-containing lentiviral vector and nonlethal irradiation. Hum Gene Ther 2006; 17: 303-313.

142 Duquette ML, Zhu Q, Taylor ER, Tsay AJ, Shi LZ, Berns MW *et al*. CtIP is required to initiate replication-dependent interstrand crosslink repair. PLoS Genet 2012; 8: e1003050.

143 Dutertre S, Ababou M, Onclercq R, Delic J, Chatton B, Jaulin C *et al*. Cell cycle regulation of the endogenous wild type Bloom's syndrome DNA helicase. Oncogene 2000; 19: 2731-2738.

144 Dutheil N, Yoon-Robarts M, Ward P, Henckaerts E, Skrabanek L, Berns KI *et al*. Characterization of the mouse adeno-associated virus AAVS1 ortholog. J Virol 2004; 78: 8917-8921.

145 Enoiu M, Jiricny J, Scharer OD. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis. Nucleic Acids Res 2012; 40: 8953-8964.

146 Epanchintsev A, Shyamsunder P, Verma RS, Lyakhovich A. IL-6, IL-8, MMP-2, MMP-9 are overexpressed in Fanconi anemia cells through a NF-kappaB/TNF-alpha dependent mechanism. Mol Carcinog 2015; 54: 1686-1699.

147 Epinat JC, Arnould S, Chames P, Rochaix P, Desfontaines D, Puzin C *et al*. A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. Nucleic Acids Res 2003; 31: 2952-2962.

148 Esashi F, Christ N, Gannon J, Liu Y, Hunt T, Jasin M *et al*. CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair. Nature 2005; 434: 598-604.

149 Escribano-Diaz C, Orthwein A, Fradet-Turcotte A, Xing M, Young JT, Tkac J *et al*. A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. Mol Cell 2013; 49: 872-883.

150 Evers B, Jonkers J. Mouse models of BRCA1 and BRCA2 deficiency: past lessons, current understanding and future prospects. Oncogene 2006; 25: 5885-5897.

151 Fagerlund RD, Staals RH, Fineran PC. The Cpf1 CRISPR-Cas protein expands genome-editing tools. Genome Biol 2015; 16: 251.

152 Fanconi anaemia/Breast cancer c. Positional cloning of the Fanconi anaemia group A gene. Nat Genet 1996; 14: 324-328.

153 Fattah FJ, Lichter NF, Fattah KR, Oh S, Hendrickson EA. Ku70, an essential gene, modulates the frequency of rAAV-mediated gene targeting in human somatic cells. Proc Natl Acad Sci U S A 2008; 105: 8703-8708.

154 Fineran PC, Charpentier E. Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information. Virology 2012; 434: 202-209.

155 Fischer A, Hacein-Bey-Abina S, Cavazzana-Calvo M. 20 years of gene therapy for SCID. Nat Immunol 2010; 11: 457-460.

156 Fischer A, Hacein-Bey Abina S, Touzot F, Cavazzana M. Gene therapy for primary immunodeficiencies. Clin Genet 2015; 88: 507-515.

157 Fishman-Lobell J, Rudin N, Haber JE. Two alternative pathways of double-strand break repair that are kinetically separable and independently modulated. Mol Cell Biol 1992; 12: 1292-1303.

158 Flygare J, Benson F, Hellgren D. Expression of the human RAD51 gene during the cell cycle in primary human peripheral blood lymphocytes. Biochim Biophys Acta 1996; 1312: 231-236.

159 Folias A, Matkovic M, Bruun D, Reid S, Hejna J, Grompe M *et al*. BRCA1 interacts directly with the Fanconi anemia protein FANCA. Hum Mol Genet 2002; 11: 2591-2597.

160 Frankenberg-Schwager M, Gebauer A, Koppe C, Wolf H, Pralle E, Frankenberg D. Single-strand annealing, conservative homologous recombination, nonhomologous DNA end joining, and the cell cycle-dependent repair of DNA double-strand breaks induced by sparsely or densely ionizing radiation. Radiation research 2009; 171: 265-273.

161 Freie B, Li X, Ciccone SL, Nawa K, Cooper S, Vogelweid C *et al*. Fanconi anemia type C and p53 cooperate in apoptosis and tumorigenesis. Blood 2003; 102: 4146-4152.

162 Friedland AE, Baral R, Singhal P, Loveluck K, Shen S, Sanchez M *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; 16: 257.

163 Frittoli MC, Biral E, Cappelli B, Zambelli M, Roncarolo MG, Ferrari G *et al*. Bone marrow as a source of hematopoietic stem cells for human gene therapy of beta-thalassemia. Hum Gene Ther 2011; 22: 507-513.

164 Fu C, Begum K, Overbeek PA. Primary Ovarian Insufficiency Induced by Fanconi Anemia E Mutation in a Mouse Model. PLoS One 2016; 11: e0144285.

165 Fu D, Dudimah FD, Zhang J, Pickering A, Paneerselvam J, Palrasu M *et al*. Recruitment of DNA polymerase eta by FANCD2 in the early response to DNA damage. Cell Cycle 2013; 12: 803-809.

166 Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol 2014; 32: 279-284.

167 Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppel C *et al*. An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol 2011; 29: 816-823.

168 Gaj T, Guo J, Kato Y, Sirk SJ, Barbas CF, 3rd. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat Methods 2012; 9: 805-807.

169 Gaj T, Epstein BE, Schaffer DV. Genome Engineering Using Adeno-associated Virus: Basic and Clinical Research Applications. Mol Ther 2016; 24: 458-464.

170 Galimi F, Noll M, Kanazawa Y, Lax T, Chen C, Grompe M *et al*. Gene therapy of Fanconi anemia: preclinical efficacy using lentiviral vectors. Blood 2002; 100: 2732-2736.

171 Gan GN, Wittschieben JP, Wittschieben BO, Wood RD. DNA polymerase zeta (pol zeta) in higher eukaryotes. Cell Res 2008; 18: 174-183.

172 Garate Z, Davis BR, Quintana-Bustamante O, Segovia JC. New frontier in regenerative medicine: site-specific gene correction in patient-specific induced pluripotent stem cells. Hum Gene Ther 2012; 24: 571-583.

173 Garaycoechea JI, Crossan GP, Langevin F, Daly M, Arends MJ, Patel KJ. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. Nature 2012; 489: 571-575.

174 Garaycoechea JI, Patel KJ. Why does the bone marrow fail in Fanconi anemia? Blood 2014; 123: 26-34.

175 Garcia-Higuera I, Kuang Y, Denham J, D'Andrea AD. The fanconi anemia proteins FANCA and FANCG stabilize each other and promote the nuclear accumulation of the Fanconi anemia complex. Blood 2000; 96: 3224-3230.

176 Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J *et al*. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. Molecular cell 2001; 7: 249-262.

177 Gari K, Decaillet C, Stasiak AZ, Stasiak A, Constantinou A. The Fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks. Molecular cell 2008; 29: 141-148.

178 Garner E, Smogorzewska A. Ubiquitylation and the Fanconi anemia pathway. FEBS Lett 2011; 585: 2853-2860.

179 Gaspar HB, Cooray S, Gilmour KC, Parsley KL, Adams S, Howe SJ *et al*. Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. Science translational medicine 2011; 3: 97ra79.

180 Gaspar HB, Cooray S, Gilmour KC, Parsley KL, Zhang F, Adams S *et al*. Hematopoietic stem cell gene therapy for adenosine deaminase-deficient severe combined immunodeficiency leads to long-term immunological recovery and metabolic correction. Science translational medicine 2011; 3: 97ra80.

181 Gaspar V, de Melo-Diogo D, Costa E, Moreira A, Queiroz J, Pichon C *et al*. Minicircle DNA vectors for gene therapy: advances and applications. Expert Opin Biol Ther 2015; 15: 353-379.

182 Gavande NS, VanderVere-Carozza PS, Hinshaw HD, Jalal SI, Sears CR, Pawelczak KS *et al*. DNA repair targeted therapy: The past or future of cancer treatment? Pharmacol Ther 2016; 160: 65-83.

183 Geng L, Huntoon CJ, Karnitz LM. RAD18-mediated ubiquitination of PCNA activates the Fanconi anemia DNA repair network. J Cell Biol 2010; 191: 249-257.

184 Genovese P, Schiroli G, Escobar G, Di Tomaso T, Firrito C, Calabria A *et al*. Targeted genome editing in human repopulating haematopoietic stem cells. Nature 2014; 510: 235-240.

185 Ghosh S, Thrasher AJ, Gaspar HB. Gene therapy for monogenic disorders of the bone marrow. Br J Haematol 2015.

186 Gill S, June CH. Going viral: chimeric antigen receptor T-cell therapy for hematological malignancies. Immunol Rev 2015; 263: 68-89.

187 Gillet LC, Scharer OD. Molecular mechanisms of mammalian global genome nucleotide excision repair. Chem Rev 2006; 106: 253-276.

188 Goebel WS, Dinauer MC. Gene therapy for chronic granulomatous disease. Acta haematologica 2003; 110: 86-92.

189 Gonzalez-Murillo A, Lozano ML, Alvarez L, Jacome A, Almarza E, Navarro S *et al*. Development of lentiviral vectors with optimized transcriptional activity for the gene therapy of patients with Fanconi anemia. Hum Gene Ther 2010; 21: 623-630.

190 Goodwin T, Huang L. Nonviral vectors: we have come a long way. Adv Genet 2014; 88: 1-12.

191 Gregory JJ, Jr., Wagner JE, Verlander PC, Levran O, Batish SD, Eide CR *et al*. Somatic mosaicism in Fanconi anemia: evidence of genotypic reversion in lymphohematopoietic stem cells. Proc Natl Acad Sci U S A 2001; 98: 2532-2537.

192 Grewal SS, Kahn JP, MacMillan ML, Ramsay NK, Wagner JE. Successful hematopoietic stem cell transplantation for Fanconi anemia from an unaffected HLA-genotype-identical sibling selected using preimplantation genetic diagnosis. Blood 2004; 103: 1147-1151.

193 Grez M, Reichenbach J, Schwable J, Seger R, Dinauer MC, Thrasher AJ. Gene therapy of chronic granulomatous disease: the engraftment dilemma. Mol Ther 2011; 19: 28-35.

194 Gross M, Hanenberg H, Lobitz S, Friedl R, Herterich S, Dietrich R *et al*. Reverse mosaicism in Fanconi anemia: natural gene therapy via molecular self-correction. Cytogenet Genome Res 2002; 98: 126-135.

195 Guardiola P, Socie G, Li X, Ribaud P, Devergie A, Esperou H *et al*. Acute graft-versus-host disease in patients with Fanconi anemia or acquired aplastic anemia undergoing bone marrow transplantation from HLA-identical sibling donors: risk factors and influence on outcome. Blood 2004; 103: 73-77.

196 Guinan EC, Lopez KD, Huhn RD, Felser JM, Nathan DG. Evaluation of granulocyte-macrophage colony-stimulating factor for treatment of pancytopenia in children with fanconi anemia. J Pediatr 1994; 124: 144-150.

197 Hacein-Bey-Abina S, Le Deist F, Carlier F, Bouneaud C, Hue C, De Villartay JP *et al*. Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. N Engl J Med 2002; 346: 1185-1193.

198 Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P *et al*. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science 2003; 302: 415-419.

199 Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E *et al*. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest 2008; 118: 3132-3142.

200 Hacein-Bey-Abina S, Hauer J, Lim A, Picard C, Wang GP, Berry CC *et al*. Efficacy of gene therapy for X-linked severe combined immunodeficiency. N Engl J Med 2010; 363: 355-364.

201 Hacein-Bey-Abina S, Pai SY, Gaspar HB, Armant M, Berry CC, Blanche S *et al*. A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency. N Engl J Med 2014; 371: 1407-1417.

202 Hacein-Bey Abina S, Gaspar HB, Blondeau J, Caccavelli L, Charrier S, Buckland K *et al*. Outcomes following gene therapy in patients with severe Wiskott-Aldrich syndrome. JAMA 2015; 313: 1550-1563.

203 Hadjur S, Ung K, Wadsworth L, Dimmick J, Rajcan-Separovic E, Scott RW *et al*. Defective hematopoiesis and hepatic steatosis in mice with combined deficiencies of the genes encoding Fancc and Cu/Zn superoxide dismutase. Blood 2001; 98: 1003-1011.

204 Handel EM, Gellhaus K, Khan K, Bednarski C, Cornu TI, Muller-Lerch F *et al*. Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adeno-associated viral vectors. Hum Gene Ther 2012; 23: 321-329.

205 Haneline LS, Gobbett TA, Ramani R, Carreau M, Buchwald M, Yoder MC *et al*. Loss of FancC function results in decreased hematopoietic stem cell repopulating ability. Blood 1999; 94: 1-8.

206 Haynes B, Saadat N, Myung B, Shekhar MP. Crosstalk between translesion synthesis, Fanconi anemia network, and homologous recombination repair pathways in interstrand DNA crosslink repair and development of chemoresistance. Mutat Res Rev Mutat Res 2015; 763: 258-266.

207 He X, Tan C, Wang F, Wang Y, Zhou R, Cui D *et al*. Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. Nucleic Acids Res 2016.

208 Helleday T, Petermann E, Lundin C, Hodgson B, Sharma RA. DNA repair pathways as targets for cancer therapy. Nat Rev Cancer 2008; 8: 193-204.

209 Heller R, Shirley S, Guo S, Donate A, Heller L. Electroporation based gene therapy--from the bench to the bedside. Conf Proc IEEE Eng Med Biol Soc 2011; 2011: 736-738.

210 Hematti P, Hong BK, Ferguson C, Adler R, Hanawa H, Sellers S *et al*. Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. PLoS biology 2004; 2: e423.

211 Henckaerts E, Dutheil N, Zeltner N, Kattman S, Kohlbrenner E, Ward P *et al*. Site-specific integration of adeno-associated virus involves partial duplication of the target locus. Proc Natl Acad Sci U S A 2009; 106: 7571-7576.

212 Henckaerts E, Linden RM. Adeno-associated virus: a key to the human genome? Future Virol 2010; 5: 555-574.

213 Hendel A, Fine EJ, Bao G, Porteus MH. Quantifying on- and off-target genome editing. Trends Biotechnol 2015; 33: 132-140.

214 Herweijer H, Wolff JA. Gene therapy progress and prospects: hydrodynamic gene delivery. Gene Ther 2007; 14: 99-107.

215 Hira A, Yoshida K, Sato K, Okuno Y, Shiraishi Y, Chiba K *et al*. Mutations in the gene encoding the E2 conjugating enzyme UBE2T cause Fanconi anemia. American journal of human genetics 2015; 96: 1001-1007.

216 Hoban MD, Cost GJ, Mendel MC, Romero Z, Kaufman ML, Joglekar AV *et al*. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. Blood 2015; 125: 2597-2604.

217 Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC *et al*. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol 2009; 27: 851-857.

218 Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP *et al*. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 2011; 29: 731-734.

219 Hodson C, Cole AR, Lewis LP, Miles JA, Purkiss A, Walden H. Structural analysis of human FANCL, the E3 ligase in the Fanconi anemia pathway. J Biol Chem 2011; 286: 32628-32637.

220 Holkers M, Maggio I, Liu J, Janssen JM, Miselli F, Mussolino C *et al*. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. Nucleic Acids Res 2013; 41: e63.

221 Holt N, Wang J, Kim K, Friedman G, Wang X, Taupin V *et al*. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol 2010; 28: 839-847.

222 Horwitz ME, Chute JP, Gasparetto C, Long GD, McDonald C, Morris A *et al*. Preemptive dosing of plerixafor given to poor stem cell mobilizers on day 5 of G-CSF administration. Bone Marrow Transplant 2012; 47: 1051-1055.

223 Houghtaling S, Timmers C, Noll M, Finegold MJ, Jones SN, Meyn MS *et al*. Epithelial cancer in Fanconi anemia complementation group D2 (Fancd2) knockout mice. Genes Dev 2003; 17: 2021-2035.

224 Houghtaling S, Granville L, Akkari Y, Torimaru Y, Olson S, Finegold M *et al*. Heterozygosity for p53 (Trp53+/-) accelerates epithelial tumor formation in fanconi anemia complementation group D2 (Fancd2) knockout mice. Cancer Res 2005; 65: 85-91.

225 Howe SJ, Mansour MR, Schwarzwaelder K, Bartholomae C, Hubank M, Kempski H *et al*. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. J Clin Invest 2008; 118: 3143-3150.

226 Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, De Die-Smulders C *et al*. Biallelic inactivation of BRCA2 in Fanconi anemia. Science 2002; 297: 606-609.

227 Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V *et al*. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 2013; 31: 827-832.

228 Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell 2014; 157: 1262-1278.

229 Hu JH, Davis KM, Liu DR. Chemical Biology Approaches to Genome Editing: Understanding, Controlling, and Delivering Programmable Nucleases. Cell Chem Biol 2016; 23: 57-73.

230 Huang J, Liu S, Bellani MA, Thazhathveetil AK, Ling C, de Winter JP *et al*. The DNA translocase FANCM/MHF promotes replication traverse of DNA interstrand crosslinks. Molecular cell 2013; 52: 434-446.

231 Huang M, Kim JM, Shiotani B, Yang K, Zou L, D'Andrea AD. The FANCM/FAAP24 complex is required for the DNA interstrand crosslink-induced checkpoint response. Molecular cell 2010; 39: 259-268.

232 Huang M, Kennedy R, Ali AM, Moreau LA, Meetei AR, D'Andrea AD *et al*. Human MutS and FANCM complexes function as redundant DNA damage sensors in the Fanconi Anemia pathway. DNA Repair (Amst) 2011; 10: 1203-1212.

233 Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B. Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 2011; 29: 699-700.

234 Huang TT, Nijman SM, Mirchandani KD, Galardy PJ, Cohn MA, Haas W *et al*. Regulation of monoubiquitinated PCNA by DUB autocleavage. Nat Cell Biol 2006; 8: 339-347.

235 Huang X, Wang Y, Yan W, Smith C, Ye Z, Wang J *et al*. Production of Gene-Corrected Adult Beta Globin Protein in Human Erythrocytes Differentiated from Patient iPSCs After Genome Editing of the Sickle Point Mutation. Stem Cells 2015; 33: 1470-1479.

236 Huang Y, Leung JW, Lowery M, Matsushita N, Wang Y, Shen X *et al*. Modularized functions of the Fanconi anemia core complex. Cell Rep 2014; 7: 1849-1857.

237 Huertas P, Jackson SP. Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. J Biol Chem 2009; 284: 9558-9565.

238 Hussain S, Witt E, Huber PA, Medhurst AL, Ashworth A, Mathew CG. Direct interaction of the Fanconi anaemia protein FANCG with BRCA2/FANCD1. Hum Mol Genet 2003; 12: 2503-2510.

239 Ibanez A, Rio P, Casado JA, Bueren JA, Fernandez-Luna JL, Pipaon C. Elevated levels of IL-1beta in Fanconi anaemia group A patients due to a constitutively active phosphoinositide 3-kinase-Akt pathway are capable of promoting tumour cell proliferation. Biochem J 2009; 422: 161-170.

240 Ikuta K, Weissman IL. Evidence that hematopoietic stem cells express mouse c-kit but do not depend on steel factor for their generation. Proc Natl Acad Sci U S A 1992; 89: 1502-1506.

241 Iliakis G, Murmann T, Soni A. Alternative end-joining repair pathways are the ultimate backup for abrogated classical non-homologous end-joining and homologous recombination repair: Implications for the formation of chromosome translocations. Mutat Res Genet Toxicol Environ Mutagen 2015; 793: 166-175.

242 Irion S, Luche H, Gadue P, Fehling HJ, Kennedy M, Keller G. Identification and targeting of the ROSA26 locus in human embryonic stem cells. Nat Biotechnol 2007; 25: 1477-1482.

243 Ishiai M, Kitao H, Smogorzewska A, Tomida J, Kinomura A, Uchida E *et al*. FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. Nat Struct Mol Biol 2008; 15: 1138-1146.

244 Ito M, Nakano T, Erdodi F, Hartshorne DJ. Myosin phosphatase: structure, regulation and function. Mol Cell Biochem 2004; 259: 197-209.

245 Ivics Z, Izsvak Z. The expanding universe of transposon technologies for gene and cell engineering. Mob DNA 2010; 1: 25.

246 Jacks T, Remington L, Williams BO, Schmitt EM, Halachmi S, Bronson RT *et al*. Tumor spectrum analysis in p53-mutant mice. Curr Biol 1994; 4: 1-7.

247 Jacome A, Navarro S, Rio P, Yanez RM, Gonzalez-Murillo A, Lozano ML *et al*. Lentiviral-mediated genetic correction of hematopoietic and mesenchymal progenitor cells from Fanconi anemia patients. Mol Ther 2009; 17: 1083-1092.

248 Joenje H, Arwert F, Eriksson AW, de Koning H, Oostra AB. Oxygen-dependence of chromosomal aberrations in Fanconi's anaemia. Nature 1981; 290: 142-143.

249 Joksic I, Vujic D, Guc-Scekic M, Leskovac A, Petrovic S, Ojani M *et al*. Dysfunctional telomeres in primary cells from Fanconi anemia FANCD2 patients. Genome Integr 2012; 3: 6.

250 Jones MJ, Huang TT. The Fanconi anemia pathway in replication stress and DNA crosslink repair. Cell Mol Life Sci 2012; 69: 3963-3974.

251 Jonkers J, Meuwissen R, van der Gulden H, Peterse H, van der Valk M, Berns A. Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. Nat Genet 2001; 29: 418-425.

252 Kahraman S, Beyazyurek C, Yesilipek MA, Ozturk G, Ertem M, Anak S *et al*. Successful haematopoietic stem cell transplantation in 44 children from healthy siblings conceived after preimplantation HLA matching. Reprod Biomed Online 2014; 29: 340-351.

253 Kang EM, Choi U, Theobald N, Linton G, Long Priel DA, Kuhns D *et al*. Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils. Blood 2010; 115: 783-791.

254 Kang HJ, Bartholomae CC, Paruzynski A, Arens A, Kim S, Yu SS *et al*. Retroviral gene therapy for X-linked chronic granulomatous disease: results from phase I/II trial. Mol Ther 2011; 19: 2092-2101.

255 Karran P. DNA double strand break repair in mammalian cells. Curr Opin Genet Dev 2000; 10: 144-150.

256 Kato Y, Alavattam KG, Sin HS, Meetei AR, Pang Q, Andreassen PR *et al*. FANCB is essential in the male germline and regulates H3K9 methylation on the sex chromosomes during meiosis. Hum Mol Genet 2015; 24: 5234-5249.

257 Kaufmann KB, Buning H, Galy A, Schambach A, Grez M. Gene therapy on the move. EMBO molecular medicine 2013; 5: 1642-1661.

258 Kay MA, Glorioso JC, Naldini L. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. Nat Med 2001; 7: 33-40.

259 Kay MA, He CY, Chen ZY. A robust system for production of minicircle DNA vectors. Nat Biotechnol 2010; 28: 1287-1289.

260 Kay MA. State-of-the-art gene-based therapies: the road ahead. Nat Rev Genet 2011; 12: 316-328.

261 Kelly PF, Radtke S, von Kalle C, Balcik B, Bohn K, Mueller R *et al*. Stem cell collection and gene transfer in Fanconi anemia. Mol Ther 2007; 15: 211-219.

262 Khalaj M, Abbasi A, Yamanishi H, Akiyama K, Wakitani S, Kikuchi S *et al*. A missense mutation in Rev7 disrupts formation of Polzeta, impairing mouse development and repair of genotoxic agent-induced DNA lesions. J Biol Chem 2014; 289: 3811-3824.

263 Khanna KK, Jackson SP. DNA double-strand breaks: signaling, repair and the cancer connection. Nat Genet 2001; 27: 247-254.

264 Kim H, D'Andrea AD. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. Genes Dev 2012; 26: 1393-1408.

265 Kim H, Kim JS. A guide to genome engineering with programmable nucleases. Nat Rev Genet 2014; 15: 321-334.

266 Kim H, Yang K, Dejsuphong D, D'Andrea AD. Regulation of Rev1 by the Fanconi anemia core complex. Nat Struct Mol Biol 2015; 19: 164-170.

267 Kim JM, Kee Y, Gurtan A, D'Andrea AD. Cell cycle-dependent chromatin loading of the Fanconi anemia core complex by FANCM/FAAP24. Blood 2008; 111: 5215-5222.

268 Kim JM, Parmar K, Huang M, Weinstock DM, Ruit CA, Kutok JL *et al*. Inactivation of murine Usp1 results in genomic instability and a Fanconi anemia phenotype. Dev Cell 2009; 16: 314-320.

269 Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome research 2014; 24: 1012-1019.

270 Kim Y, Lach FP, Desetty R, Hanenberg H, Auerbach AD, Smogorzewska A. Mutations of the SLX4 gene in Fanconi anemia. Nat Genet 2011; 43: 142-146.

271 Kim YG, Chandrasegaran S. Chimeric restriction endonuclease. Proc Natl Acad Sci U S A 1994; 91: 883-887.

272 Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 1996; 93: 1156-1160.

273 Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z *et al*. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature 2016; 529: 490-495.

274 Kohn DB. Update on gene therapy for immunodeficiencies. Clin Immunol 2010; 135: 247-254.

275 Koo T, Lee J, Kim JS. Measuring and Reducing Off-Target Activities of Programmable Nucleases Including CRISPR-Cas9. Mol Cells 2015; 38: 475-481.

276 Koomen M, Cheng NC, van de Vrugt HJ, Godthelp BC, van der Valk MA, Oostra AB *et al*. Reduced fertility and hypersensitivity to mitomycin C characterize Fancg/Xrcc9 null mice. Hum Mol Genet 2002; 11: 273-281.

277 Kotin RM, Linden RM, Berns KI. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. EMBO J 1992; 11: 5071-5078.

278 Kottemann MC, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. Nature 2013; 493: 356-363.

279 Kowal P, Gurtan AM, Stuckert P, D'Andrea AD, Ellenberger T. Structural determinants of human FANCF protein that function in the assembly of a DNA damage signaling complex. J Biol Chem 2007; 282: 2047-2055.

280 Kuhn AN, Beibetaert T, Simon P, Vallazza B, Buck J, Davies BP *et al*. mRNA as a versatile tool for exogenous protein expression. Current gene therapy 2012; 12: 347-361.

281 Kumaraswamy E, Shiekhattar R. Activation of BRCA1/BRCA2-associated helicase BACH1 is required for timely progression through S phase. Mol Cell Biol 2007; 27: 6733-6741.

282 Kung SH, Retchless AC, Kwan JY, Almeida RP. Effects of DNA size on transformation and recombination efficiencies in Xylella fastidiosa. Appl Environ Microbiol 2013; 79: 1712-1717.

283 Kuo CY, Kohn DB. Gene Therapy for the Treatment of Primary Immune Deficiencies. Curr Allergy Asthma Rep 2016; 16: 39.

284 Kuznetsov S, Pellegrini M, Shuda K, Fernandez-Capetillo O, Liu Y, Martin BK *et al*. RAD51C deficiency in mice results in early prophase I arrest in males and sister chromatid separation at metaphase II in females. J Cell Biol 2007; 176: 581-592.

285 Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970; 227: 680-685.

286 Lambricht L, Lopes A, Kos S, Sersa G, Preat V, Vandermeulen G. Clinical potential of electroporation for gene therapy and DNA vaccine delivery. Expert Opin Drug Deliv 2015; 13: 295-310.

287 Langevin F, Crossan GP, Rosado IV, Arends MJ, Patel KJ. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. Nature 2011; 475: 53-58.

288 Langevin FP, Garaycoechea JI, Crossan GP, Patel KJ. [Aldehydes and Fanconi anaemia: the enemy within]. Med Sci (Paris) 2013; 29: 361-364.

289 Larghero J, Marolleau JP, Soulier J, Filion A, Rocha V, Benbunan M *et al*. Hematopoietic progenitor cel harvest and functionality in Fanconi anemia patients. Blood 2002; 100: 3051.

290 Leskovac A, Petrovic S, Guc-Scekic M, Vujic D, Joksic G. Radiation-induced mitotic catastrophe in FANCD2 primary fibroblasts. Int J Radiat Biol 2014; 90: 373-381.

291 Leteurtre F, Li X, Guardiola P, Le Roux G, Sergere JC, Richard P *et al*. Accelerated telomere shortening and telomerase activation in Fanconi's anaemia. Br J Haematol 1999; 105: 883-893.

292 Leung JW, Wang Y, Fong KW, Huen MS, Li L, Chen J. Fanconi anemia (FA) binding protein FAAP20 stabilizes FA complementation group A (FANCA) and participates in interstrand cross-link repair. Proc Natl Acad Sci U S A 2012; 109: 4491-4496.

293 Levitus M, Rooimans MA, Steltenpool J, Cool NF, Oostra AB, Mathew CG *et al*. Heterogeneity in Fanconi anemia: evidence for 2 new genetic subtypes. Blood 2004; 103: 2498-2503.

294 Levitus M, Waisfisz Q, Godthelp BC, de Vries Y, Hussain S, Wiegant WW *et al*. The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J. Nat Genet 2005; 37: 934-935.

295 Li HL, Fujimoto N, Sasakawa N, Shirai S, Ohkame T, Sakuma T *et al*. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. Stem cell reports 2015; 4: 143-154.

296 Li J, Sejas DP, Zhang X, Qiu Y, Nattamai KJ, Rani R *et al*. TNF-alpha induces leukemic clonal evolution ex vivo in Fanconi anemia group C murine stem cells. J Clin Invest 2007; 117: 3283-3295.

297 Li L, Krymskaya L, Wang J, Henley J, Rao A, Cao LF *et al*. Genomic editing of the HIV-1 coreceptor CCR5 in adult hematopoietic stem and progenitor cells using zinc finger nucleases. Mol Ther 2013; 21: 1259-1269.

298 Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP *et al*. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res 2011; 39: 359-372.

299 Li X, Heyer WD. RAD54 controls access to the invading 3'-OH end after RAD51-mediated DNA strand invasion in homologous recombination in Saccharomyces cerevisiae. Nucleic Acids Res 2009; 37: 638-646.

300 Lieber MR, Ma Y, Pannicke U, Schwarz K. Mechanism and regulation of human non-homologous DNA end-joining. Nat Rev Mol Cell Biol 2003; 4: 712-720.

301 Lim DS, Hasty P. A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. Mol Cell Biol 1996; 16: 7133-7143.

302 Lim ET, Wurtz P, Havulinna AS, Palta P, Tukiainen T, Rehnstrom K *et al*. Distribution and medical impact of loss-of-function variants in the Finnish founder population. PLoS Genet 2014; 10: e1004494.

303 Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Elife 2014; 3: e04766.

304 Linden RM, Ward P, Giraud C, Winocour E, Berns KI. Site-specific integration by adeno-associated virus. Proc Natl Acad Sci U S A 1996; 93: 11288-11294.

305 Linden RM, Winocour E, Berns KI. The recombination signals for adeno-associated virus site-specific integration. Proc Natl Acad Sci U S A 1996; 93: 7966-7972.

306 Ling C, Ishiai M, Ali AM, Medhurst AL, Neveling K, Kalb R *et al*. FAAP100 is essential for activation of the Fanconi anemia-associated DNA damage response pathway. EMBO J 2007; 26: 2104-2114.

307 Liu F, Song Y, Liu D. Hydrodynamics-based transfection in animals by systemic administration of plasmid DNA. Gene Ther 1999; 6: 1258-1266.

308 Liu F, Tyagi P. Naked DNA for liver gene transfer. Adv Genet 2005; 54: 43-64.

309 Liu GH, Suzuki K, Li M, Qu J, Montserrat N, Tarantino C *et al*. Modelling Fanconi anemia pathogenesis and therapeutics using integration-free patient-derived iPSCs. Nat Commun 2014; 5: 4330.

310 Liu J, Gaj T, Patterson JT, Sirk SJ, Barbas CF, 3rd. Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. PLoS One 2014; 9: e85755.

311 Liu JM, Kim S, Read EJ, Futaki M, Dokal I, Carter CS *et al*. Engraftment of hematopoietic progenitor cells transduced with the Fanconi anemia group C gene (FANCC). Hum Gene Ther 1999; 10: 2337-2346.

312 Liu Z, Venkatesh SS, Maley CC. Sequence space coverage, entropy of genomes and the potential to detect non-human DNA in human samples. BMC genomics 2008; 9: 509.

313 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001; 25: 402-408.

314 Lo Ten Foe JR, Rooimans MA, Bosnoyan-Collins L, Alon N, Wijker M, Parker L *et al*. Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA. Nat Genet 1996; 14: 320-323.

315 Lo Ten Foe JR, Kwee ML, Rooimans MA, Oostra AB, Veerman AJ, van Weel M *et al*. Somatic mosaicism in Fanconi anemia: molecular basis and clinical significance. Eur J Hum Genet 1997; 5: 137-148.

316 Lobitz S, Velleuer E. Guido Fanconi (1892-1979): a jack of all trades. Nat Rev Cancer 2006; 6: 893-898.

317 Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA *et al*. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol 2007; 25: 1298-1306.

318 Lombardo A, Cesana D, Genovese P, Di Stefano B, Provasi E, Colombo DF *et al*. Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nat Methods 2011; 8: 861-869.

319 Longerich S, Kwon Y, Tsai MS, Hlaing AS, Kupfer GM, Sung P. Regulation of FANCD2 and FANCI monoubiquitination by their interaction and by DNA. Nucleic Acids Res 2014; 42: 5657-5670.

320 Longerich S, Li J, Xiong Y, Sung P, Kupfer GM. Stress and DNA repair biology of the Fanconi anemia pathway. Blood 2014; 124: 2812-2819.

321 Ma N, Liao B, Zhang H, Wang L, Shan Y, Xue Y *et al*. Transcription activator-like effector nuclease (TALEN)-mediated gene correction in integration-free beta-thalassemia induced pluripotent stem cells. J Biol Chem 2013; 288: 34671-34679.

322 Ma N, Shan Y, Liao B, Kong G, Wang C, Huang K *et al*. Factor-induced Reprogramming and Zinc Finger Nuclease-aided Gene Targeting Cause Different Genome Instability in beta-Thalassemia Induced Pluripotent Stem Cells (iPSCs). J Biol Chem 2015; 290: 12079-12089.

323 Macmillan ML, Wagner JE. Haematopoeitic cell transplantation for Fanconi anaemia - when and how? Br J Haematol 2010; 149: 14.21.

324 MacMillan ML, DeFor TE, Young JA, Dusenbery KE, Blazar BR, Slungaard A *et al*. Alternative donor hematopoietic cell transplantation for Fanconi anemia. Blood 2015; 125: 3798-3804.

325 Maeder ML, Gersbach CA. Genome-editing Technologies for Gene and Cell Therapy. Mol Ther 2016; 24: 430-446.

326 Maetzig T, Galla M, Baum C, Schambach A. Gammaretroviral vectors: biology, technology and application. Viruses 2011; 3: 677-713.

327 Maggio I, Goncalves MA. Genome editing at the crossroads of delivery, specificity, and fidelity. Trends Biotechnol 2015; 33: 280-291.

328 Mahfouz MM, Li L, Shamimuzzaman M, Wibowo A, Fang X, Zhu JK. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci U S A 2011; 108: 2623-2628.

329 Malech HL, Maples PB, Whiting-Theobald N, Linton GF, Sekhsaria S, Vowells SJ *et al*. Prolonged production of NADPH oxidase-corrected granulocytes after gene therapy of chronic granulomatous disease. Proc Natl Acad Sci U S A 1997; 94: 12133-12138.

330 Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE *et al*. RNA-guided human genome engineering via Cas9. Science 2013; 339: 823-826.

331 Mankad A, Taniguchi T, Cox B, Akkari Y, Rathbun RK, Lucas L *et al*. Natural gene therapy in monozygotic twins with Fanconi anemia. Blood 2006; 107: 3084-3090.

332 Mansour WY, Schumacher S, Rosskopf R, Rhein T, Schmidt-Petersen F, Gatzemeier F *et al*. Hierarchy of nonhomologous end-joining, single-strand annealing and gene conversion at site-directed DNA double-strand breaks. Nucleic Acids Res 2008; 36: 4088-4098.

333 Manthey GM, Bailis AM. Rad51 inhibits translocation formation by non-conservative homologous recombination in Saccharomyces cerevisiae. PLoS One 2010; 5: e11889.

334 Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol 2015; 33: 538-542.

335 Masat E, Pavani G, Mingozzi F. Humoral immunity to AAV vectors in gene therapy: challenges and potential solutions. Discov Med 2013; 15: 379-389.

336 Mason PJ, Bessler M. Cytokinesis failure and attenuation: new findings in Fanconi anemia. J Clin Invest 2011; 121: 27-30.

337 Matsubara Y, Chiba T, Kashimada K, Morio T, Takada S, Mizutani S *et al*. Transcription activator-like effector nuclease-mediated transduction of exogenous gene into IL2RG locus. Sci Rep 2014; 4: 5043.

338 Matsushita N, Endo Y, Sato K, Kurumizaka H, Yamashita T, Takata M *et al*. Direct inhibition of TNF-alpha promoter activity by Fanconi anemia protein FANCD2. PLoS One 2011; 6: e23324.

339 McAllister KA, Bennett LM, Houle CD, Ward T, Malphurs J, Collins NK *et al*. Cancer susceptibility of mice with a homozygous deletion in the COOH-terminal domain of the Brca2 gene. Cancer Res 2002; 62: 990-994.

340 Medhurst AL, Laghmani el H, Steltenpool J, Ferrer M, Fontaine C, de Groot J *et al*. Evidence for subcomplexes in the Fanconi anemia pathway. Blood 2006; 108: 2072-2080.

341 Meek K, Gupta S, Ramsden DA, Lees-Miller SP. The DNA-dependent protein kinase: the director at the end. Immunol Rev 2004; 200: 132-141.

342 Meetei AR, de Winter JP, Medhurst AL, Wallisch M, Waisfisz Q, van de Vrugt HJ *et al*. A novel ubiquitin ligase is deficient in Fanconi anemia. Nat Genet 2003; 35: 165-170.

343 Meetei AR, Levitus M, Xue Y, Medhurst AL, Zwaan M, Ling C *et al*. X-linked inheritance of Fanconi anemia complementation group B. Nat Genet 2004; 36: 1219-1224.

344 Meetei AR, Medhurst AL, Ling C, Xue Y, Singh TR, Bier P *et al*. A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. Nat Genet 2005; 37: 958-963.

345 Menon T, Firth AL, Scripture-Adams DD, Galic Z, Qualls SJ, Gilmore WB *et al*. Lymphoid regeneration from gene-corrected SCID-X1 subject-derived iPSCs. Cell Stem Cell 2015; 16: 367-372.

346 Merkert S, Martin U. Targeted genome engineering using designer nucleases: State of the art and practical guidance for application in human pluripotent stem cells. Stem Cell Res 2016; 16: 377-386.

347 Merling RK, Sweeney CL, Chu J, Bodansky A, Choi U, Priel DL *et al*. An AAVS1-targeted minigene platform for correction of iPSCs from all five types of chronic granulomatous disease. Mol Ther 2015; 23: 147-157.

348 Merten OW, Charrier S, Laroudie N, Fauchille S, Dugue C, Jenny C *et al*. Large-scale manufacture and characterization of a lentiviral vector produced for clinical ex vivo gene therapy application. Hum Gene Ther 2011; 22: 343-356.

349 Miccio A, Cesari R, Lotti F, Rossi C, Sanvito F, Ponzoni M *et al*. In vivo selection of genetically modified erythroblastic progenitors leads to long-term correction of beta-thalassemia. Proc Natl Acad Sci U S A 2008; 105: 10547-10552.

350 Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF *et al*. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 2011; 29: 143-148.

351 Mitchell RS, Beitzel BF, Schroder AR, Shinn P, Chen H, Berry CC *et al*. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS biology 2004; 2: E234.

352 Mizutani T, Li R, Haga H, Kawabata K. Transgene integration into the human AAVS1 locus enhances myosin II-dependent contractile force by reducing expression of myosin binding subunit 85. Biochem Biophys Res Commun 2015; 465: 270-274.

353 Mizutani T, Haga H, Kawabata K. Data set for comparison of cellular dynamics between human AAVS1 locus-modified and wild-type cells. Data Brief 2016; 6: 793-798.

354 Mladenov E, Magin S, Soni A, Iliakis G. DNA double-strand break repair as determinant of cellular radiosensitivity to killing and target in radiation therapy. Front Oncol 2013; 3: 113.

355 Mladenov E, Magin S, Soni A, Iliakis G. DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: Cell cycle and proliferation-dependent regulation. Semin Cancer Biol 2016.

356 Mock U, Riecken K, Berdien B, Qasim W, Chan E, Cathomen T *et al*. Novel lentiviral vectors with mutated reverse transcriptase for mRNA delivery of TALE nucleases. Sci Rep 2014; 4: 6409.

357 Mock U, Machowicz R, Hauber I, Horn S, Abramowski P, Berdien B *et al*. mRNA transfection of a novel TAL effector nuclease (TALEN) facilitates efficient knockout of HIV co-receptor CCR5. Nucleic Acids Res 2015; 43: 5560-5571.

358 Modlich U, Navarro S, Zychlinski D, Maetzig T, Knoess S, Brugman MH *et al*. Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. Mol Ther 2009; 17: 1919-1928.

359 Moehle EA, Rock JM, Lee YL, Jouvenot Y, DeKelver RC, Gregory PD *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.

360 Mohrin M, Bourke E, Alexander D, Warr MR, Barry-Holson K, Le Beau MM *et al*. Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. Cell Stem Cell 2010; 7: 174-185.

361 Moiani A, Paleari Y, Sartori D, Mezzadra R, Miccio A, Cattoglio C *et al*. Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts. J Clin Invest 2012; 122: 1653-1666.

362 Moldovan GL, Pfander B, Jentsch S. PCNA, the maestro of the replication fork. Cell 2007; 129: 665-679.

363 Moldovan GL, D'Andrea AD. How the fanconi anemia pathway guards the genome. Annu Rev Genet 2009; 43: 223-249.

364 Moldovan GL, Dejsuphong D, Petalcorin MI, Hofmann K, Takeda S, Boulton SJ *et al*. Inhibition of homologous recombination by the PCNA-interacting protein PARI. Molecular cell 2012; 45: 75-86.

365 Molina-Estevez FJ, Nowrouzi A, Lozano ML, Galy A, Charrier S, von Kalle C *et al*. Lentiviral-Mediated Gene Therapy in Fanconi Anemia-A Mice Reveals Long-Term

Engraftment and Continuous Turnover of Corrected HSCs. Current gene therapy 2015; 15: 550-562.

366 Molina B, Marchetti F, Gomez L, Ramos S, Torres L, Ortiz R *et al*. Hydroxyurea induces chromosomal damage in G2 and enhances the clastogenic effect of mitomycin C in Fanconi anemia cells. Environ Mol Mutagen 2015.

367 Montini E, Cesana D, Schmidt M, Sanvito F, Ponzoni M, Bartholomae C *et al*. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. Nat Biotechnol 2006; 24: 687-696.

368 Montini E, Cesana D, Schmidt M, Sanvito F, Bartholomae CC, Ranzani M *et al*. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. J Clin Invest 2009; 119: 964-975.

369 Morbitzer R, Elsaesser J, Hausner J, Lahaye T. Assembly of custom TALE-type DNA binding domains by modular cloning. Nucleic Acids Res 2011; 39: 5790-5799.

370 Morrison SJ, Uchida N, Weissman IL. The biology of hematopoietic stem cells. Annu Rev Cell Dev Biol 1995; 11: 35-71.

371 Mostoslavsky G, Kotton DN, Fabian AJ, Gray JT, Lee JS, Mulligan RC. Efficiency of transduction of highly purified murine hematopoietic stem cells by lentiviral and oncoretroviral vectors under conditions of minimal in vitro manipulation. Mol Ther 2005; 11: 932-940.

372 Mukherjee S, Thrasher AJ. Gene therapy for PIDs: progress, pitfalls and prospects. Gene 2013; 525: 174-181.

373 Mullen CA, Snitzer K, Culver KW, Morgan RA, Anderson WF, Blaese RM. Molecular analysis of T lymphocyte-directed gene therapy for adenosine deaminase deficiency: long-term expression in vivo of genes introduced with a retroviral vector. Hum Gene Ther 1996; 7: 1123-1129.

374 Muller LU, Milsom MD, Kim MO, Schambach A, Schuesler T, Williams DA. Rapid lentiviral transduction preserves the engraftment potential of Fanca(-/-) hematopoietic stem cells. Mol Ther 2008; 16: 1154-1160.

375 Murina O, von Aesch C, Karakus U, Ferretti LP, Bolck HA, Hanggi K *et al*. FANCD2 and CtIP cooperate to repair DNA interstrand crosslinks. Cell Rep 2014; 7: 1030-1038.

376 Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res 2011; 39: 9283-9293.

377 Mussolino C, Cathomen T. TALE nucleases: tailored genome engineering made easy. Curr Opin Biotechnol 2012; 23: 644-650.

378 Mussolino C, Alzubi J, Fine EJ, Morbitzer R, Cradick TJ, Lahaye T *et al*. TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. Nucleic Acids Res 2014; 42: 6762-6773.

379 Muul LM, Tuschong LM, Soenen SL, Jagadeesh GJ, Ramsey WJ, Long Z *et al*. Persistence and expression of the adenosine deaminase gene for 12 years and immune reaction to gene transfer components: long-term results of the first clinical gene therapy trial. Blood 2003; 101: 2563-2569.

380 Nakamura-Ishizu A, Takizawa H, Suda T. The analysis, roles and regulation of quiescence in hematopoietic stem cells. Development 2014; 141: 4656-4666.

381 Nakanishi K, Yang YG, Pierce AJ, Taniguchi T, Digweed M, D'Andrea AD *et al*. Human Fanconi anemia monoubiquitination pathway promotes homologous DNA repair. Proc Natl Acad Sci U S A 2005; 102: 1110-1115.

382 Nakanishi K, Cavallo F, Perrouault L, Giovannangeli C, Moynahan ME, Barchi M *et al*. Homology-directed Fanconi anemia pathway cross-link repair is dependent on DNA replication. Nat Struct Mol Biol 2011; 18: 500-503.

383 Naldini L. Ex vivo gene transfer and correction for cell-based therapies. Nat Rev Genet 2011; 12: 301-315.

384 Naldini L. Gene therapy returns to centre stage. Nature 2015; 526: 351-360.

385 Nalepa G, Enzor R, Sun Z, Marchal C, Park SJ, Yang Y *et al*. Fanconi anemia signaling network regulates the spindle assembly checkpoint. J Clin Invest 2013; 123: 3839-3847.

386 Navarro S, Meza NW, Quintana-Bustamante O, Casado JA, Jacome A, McAllister K *et al*. Hematopoietic dysfunction in a mouse model for Fanconi anemia group D1. Mol Ther 2006; 14: 525-535.

387 Navarro S, Rio P, Bueren J. Perspectives on gene therapy for Fanconi anemia. Exper Opinion on Orphan Durgs 2015; 3: 1-12.

388 Naviaux RK, Costanzi E, Haas M, Verma IM. The pCL vector system: rapid production of helper-free, high-titer, recombinant retroviruses. J Virol 1996; 70: 5701-5705.

389 Negre O, Eggimann AV, Beuzard Y, Ribeil JA, Bourget P, Borwornpinyo S *et al*. Gene Therapy of the beta-Hemoglobinopathies by Lentiviral Transfer of the beta(A(T87Q))-Globin Gene. Hum Gene Ther 2016; 27: 148-165.

390 Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Castellanos Rivera RM *et al*. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science 2016; 351: 403-407.

391 Neuberg P, Kichler A. Recent developments in nucleic acid delivery with polyethylenimines. Adv Genet 2014; 88: 263-288.

392 Neveling K, Endt D, Hoehn H, Schindler D. Genotype-phenotype correlations in Fanconi anemia. Mutat Res 2009; 668: 73-91.

393 Niedernhofer LJ, Garinis GA, Raams A, Lalai AS, Robinson AR, Appeldoorn E *et al*. A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. Nature 2006; 444: 1038-1043.

394 Nijman SM, Huang TT, Dirac AM, Brummelkamp TR, Kerkhoven RM, D'Andrea AD *et al*. The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway. Molecular cell 2005; 17: 331-339.

395 Nojima K, Hochegger H, Saberi A, Fukushima T, Kikuchi K, Yoshimura M *et al*. Multiple repair pathways mediate tolerance to chemotherapeutic cross-linking agents in vertebrate cells. Cancer Res 2005; 65: 11704-11711.

396 Noll M, Battaile KP, Bateman R, Lax TP, Rathbun K, Reifsteck C *et al*. Fanconi anemia group A and C double-mutant mice: functional evidence for a multi-protein Fanconi anemia complex. Exp Hematol 2002; 30: 679-688.

397 Nowakowski A, Andrzejewska A, Janowski M, Walczak P, Lukomska B. Genetic engineering of stem cells for enhanced therapy. Acta Neurobiol Exp (Wars) 2013; 73: 1-18.

398 Oceguera-Yanez F, Kim SI, Matsumoto T, Tan GW, Xiang L, Hatani T *et al*. Engineering the AAVS1 locus for consistent and scalable transgene expression in human iPSCs and their differentiated derivatives. Methods.

399 Onn I, Heidinger-Pauli JM, Guacci V, Unal E, Koshland DE. Sister chromatid cohesion: a simple concept with a complex reality. Annu Rev Cell Dev Biol 2008; 24: 105-129.

400 Ordovas L, Boon R, Pistoni M, Chen Y, Wolfs E, Guo W *et al*. Efficient Recombinase-Mediated Cassette Exchange in hPSCs to Study the Hepatocyte Lineage Reveals AAVS1 Locus-Mediated Transgene Inhibition. Stem cell reports 2015; 5: 918-931.

401 Orlando SJ, Santiago Y, DeKelver RC, Freyvert Y, Boydston EA, Moehle EA *et al*. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res 2010; 38: e152.

402 Osborn M, Lonetree CL, Webber BR, Patel D, Dunmire S, McElroy AN *et al*. CRISPR/Cas9 Targeted Gene Editing and Cellular Engineering in Fanconi Anemia. Stem Cells Dev 2016.

403 Osborn MJ, Gabriel R, Webber BR, DeFeo AP, McElroy AN, Jarjour J *et al*. Fanconi anemia gene editing by the CRISPR/Cas9 system. Hum Gene Ther 2015; 26: 114-126.

404 Ott de Bruin LM, Volpi S, Musunuru K. Novel Genome-Editing Tools to Model and Correct Primary Immunodeficiencies. Front Immunol 2015; 6: 250.

405 Ott MG, Schmidt M, Schwarzwaelder K, Stein S, Siler U, Koehl U *et al*. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat Med 2006; 12: 401-409.

406 Ousterout DG, Perez-Pinera P, Thakore PI, Kabadi AM, Brown MT, Qin X *et al*. Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. Mol Ther 2013; 21: 1718-1726.

407 Overmeer RM, Moser J, Volker M, Kool H, Tomkinson AE, van Zeeland AA *et al*. Replication protein A safeguards genome integrity by controlling NER incision events. J Cell Biol 2011; 192: 401-415.

408 Pace P, Mosedale G, Hodskinson MR, Rosado IV, Sivasubramaniam M, Patel KJ. Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. Science 2010; 329: 219-223.

409 Pagano G, Degan P, d'Ischia M, Kelly FJ, Nobili B, Pallardo FV *et al*. Oxidative stress as a multiple effector in Fanconi anaemia clinical phenotype. Eur J Haematol 2005; 75: 93-100.

410 Pagano G, Talamanca AA, Castello G, Pallardo FV, Zatterale A, Degan P. Oxidative stress in Fanconi anaemia: from cells and molecules towards prospects in clinical management. Biol Chem 2012; 393: 11-21.

411 Pagano G, d'Ischia M, Pallardo FV. Fanconi anemia (FA) and crosslinker sensitivity: Re-appraising the origins of FA definition. Pediatr Blood Cancer 2015; 62: 1137-1143.

412 Paiboonsukwong K, Ohbayashi F, Shiiba H, Aizawa E, Yamashita T, Mitani K. Correction of mutant Fanconi anemia gene by homologous recombination in human hematopoietic cells using adeno-associated virus vector. J Gene Med 2009; 11: 1012-1019.

413 Palacios R, Henson G, Steinmetz M, McKearn JP. Interleukin-3 supports growth of mouse pre-B-cell clones in vitro. Nature 1984; 309: 126-131.

414 Palacios R, Steinmetz M. Il-3-dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ-line configuration, and generate B lymphocytes in vivo. Cell 1985; 41: 727-734.

415 Pallardo FV, Lloret A, Lebel M, d'Ischia M, Cogger VC, Le Couteur DG *et al*. Mitochondrial dysfunction in some oxidative stress-related genetic diseases: Ataxia-Telangiectasia, Down Syndrome, Fanconi Anaemia and Werner Syndrome. Biogerontology 2010; 11: 401-419.

416 Papapetrou EP, Lee G, Malani N, Setty M, Riviere I, Tirunagari LM *et al*. Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. Nat Biotechnol 2011; 29: 73-78.

417 Papapetrou EP, Schambach A. Gene Insertion Into Genomic Safe Harbors for Human Gene Therapy. Mol Ther 2016; 24: 678-684.

418 Park CY, Kim J, Kweon J, Son JS, Lee JS, Yoo JE *et al*. Targeted inversion and reversion of the blood coagulation factor 8 gene in human iPS cells using TALENs. Proc Natl Acad Sci U S A 2014; 111: 9253-9258.

419 Park CY, Kim DH, Son JS, Sung JJ, Lee J, Bae S *et al*. Functional Correction of Large Factor VIII Gene Chromosomal Inversions in Hemophilia A Patient-Derived iPSCs Using CRISPR-Cas9. Cell Stem Cell 2015; 17: 213-220.

420 Parmar K, D'Andrea A, Niedernhofer LJ. Mouse models of Fanconi anemia. Mutat Res 2009; 668: 133-140.

421 Parmar K, Kim J, Sykes SM, Shimamura A, Stuckert P, Zhu K *et al*. Hematopoietic stem cell defects in mice with deficiency of Fancd2 or Usp1. Stem Cells 2010; 28: 1186-1195.

422 Pear WS, Nolan GP, Scott ML, Baltimore D. Production of high-titer helper-free retroviruses by transient transfection. Proc Natl Acad Sci U S A 1993; 90: 8392-8396.

423 Peffault de Latour R, Porcher R, Dalle JH, Aljurf M, Korthof ET, Svahn J *et al*. Allogeneic hematopoietic stem cell transplantation in Fanconi anemia: the European Group for Blood and Marrow Transplantation experience. Blood 2013; 122: 4279-4286.

424 Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O *et al*. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol 2008; 26: 808-816.

425 Petrezselyova S, Kinsky S, Truban D, Sedlacek R, Burtscher I, Lickert H. Homology arms of targeting vectors for gene insertions and CRISPR/Cas9 technology: size does not matter; quality control of targeted clones does. Cell Mol Biol Lett 2015; 20: 773-787.

426 Petruseva IO, Evdokimov AN, Lavrik OI. Molecular mechanism of global genome nucleotide excision repair. Acta Naturae 2014; 6: 23-34.

427 Pichierri P, Rosselli F. The DNA crosslink-induced S-phase checkpoint depends on ATR-CHK1 and ATR-NBS1-FANCD2 pathways. EMBO J 2004; 23: 1178-1187.

428 Pilonetto DV, Pereira NF, Bitencourt MA, Magdalena NI, Vieira ER, Veiga LB *et al*. FANCD2 Western blot as a diagnostic tool for Brazilian patients with Fanconi anemia. Braz J Med Biol Res 2009; 42: 237-243.

429 Pinder J, Salsman J, Dellaire G. Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. Nucleic Acids Res 2015; 43: 9379-9392.

430 Poot M, Gross O, Epe B, Pflaum M, Hoehn H. Cell cycle defect in connection with oxygen and iron sensitivity in Fanconi anemia lymphoblastoid cells. Exp Cell Res 1996; 222: 262-268.

431 Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. Science 2003; 300: 763.

432 Porteus MH, Connelly JP, Pruett SM. A look to future directions in gene therapy research for monogenic diseases. PLoS Genet 2006; 2: e133.

433 Prakash R, Zhang Y, Feng W, Jasin M. Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. Cold Spring Harb Perspect Biol 2015; 7: a016600.

434 Prakash V, Moore M, Yanez-Munoz RJ. Current Progress in Therapeutic Gene Editing for Monogenic Diseases. Mol Ther 2016; 24: 465-474.

435 Pulliam-Leath AC, Ciccone SL, Nalepa G, Li X, Si Y, Miravalle L *et al*. Genetic disruption of both Fancc and Fancg in mice recapitulates the hematopoietic manifestations of Fanconi anemia. Blood 2010; 116: 2915-2920.

436 Pulliam AC, Hobson MJ, Ciccone SL, Li Y, Chen S, Srour EF *et al*. AMD3100 synergizes with G-CSF to mobilize repopulating stem cells in Fanconi anemia knockout mice. Exp Hematol 2008; 36: 1084-1090.

437 Qian K, Huang CT, Chen H, Blackbourn LWt, Chen Y, Cao J *et al*. A simple and efficient system for regulating gene expression in human pluripotent stem cells and derivatives. Stem Cells 2014; 32: 1230-1238.

438 Rackoff WR, Orazi A, Robinson CA, Cooper RJ, Alter BP, Freedman MH *et al*. Prolonged administration of granulocyte colony-stimulating factor (filgrastim) to patients with Fanconi anemia: a pilot study. Blood 1996; 88: 1588-1593.

439 Radecke S, Radecke F, Cathomen T, Schwarz K. Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. Mol Ther 2010; 18: 743-753.

440 Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A *et al*. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nat Genet 2007; 39: 165-167.

441 Rahman SH, Maeder ML, Joung JK, Cathomen T. Zinc-finger nucleases for somatic gene therapy: the next frontier. Hum Gene Ther 2011; 22: 925-933.

442 Rahman SH, Bobis-Wozowicz S, Chatterjee D, Gellhaus K, Pars K, Heilbronn R *et al*. The nontoxic cell cycle modulator indirubin augments transduction of adeno-associated viral vectors and zinc-finger nuclease-mediated gene targeting. Hum Gene Ther 2013; 24: 67-77.

443 Rahman SH, Kuehle J, Reimann C, Mlambo T, Alzubi J, Maeder ML *et al*. Rescue of DNA-PK Signaling and T-Cell Differentiation by Targeted Genome Editing in a prkdc Deficient iPSC Disease Model. PLoS Genet 2015; 11: e1005239.

444 Rajendra E, Oestergaard VH, Langevin F, Wang M, Dornan GL, Patel KJ *et al*. The genetic and biochemical basis of FANCD2 monoubiquitination. Molecular cell 54: 858-869.

445 Ramachandra CJ, Shahbazi M, Kwang TW, Choudhury Y, Bak XY, Yang J *et al*. Efficient recombinase-mediated cassette exchange at the AAVS1 locus in human embryonic stem cells using baculoviral vectors. Nucleic Acids Res 2011; 39: e107.

446 Ramakrishna S, Kwaku Dad AB, Beloor J, Gopalappa R, Lee SK, Kim H. Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. Genome research 2014; 24: 1020-1027.

447 Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE *et al*. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 2013; 154: 1380-1389.

448 Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ *et al*. In vivo genome editing using Staphylococcus aureus Cas9. Nature 2015; 520: 186-191.

449 Rantakari P, Nikkila J, Jokela H, Ola R, Pylkas K, Lagerbohm H *et al*. Inactivation of Palb2 gene leads to mesoderm differentiation defect and early embryonic lethality in mice. Hum Mol Genet 2010; 19: 3021-3029.

450 Raschle M, Knipscheer P, Enoiu M, Angelov T, Sun J, Griffith JD *et al*. Mechanism of replication-coupled DNA interstrand crosslink repair. Cell 2008; 134: 969-980.

451 Reaume AG, Elliott JL, Hoffman EK, Kowall NW, Ferrante RJ, Siwek DF *et al*. Motor neurons in Cu/Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury. Nat Genet 1996; 13: 43-47.

452 Reid S, Schindler D, Hanenberg H, Barker K, Hanks S, Kalb R *et al*. Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. Nat Genet 2007; 39: 162-164.

453 Reliene R, Yamamoto ML, Rao PN, Schiestl RH. Genomic instability in mice is greater in Fanconi anemia caused by deficiency of Fancd2 than Fancg. Cancer Res 2010; 70: 9703-9710.

454 Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 2012; 30: 460-465.

455 Rio P, Segovia JC, Hanenberg H, Casado JA, Martinez J, Gottsche K *et al*. In vitro phenotypic correction of hematopoietic progenitors from Fanconi anemia group A knockout mice. Blood 2002; 100: 2032-2039.

456 Rio P, Banos R, Lombardo A, Quintana-Bustamante O, Alvarez L, Garate Z *et al*. Targeted gene therapy and cell reprogramming in Fanconi anemia. EMBO molecular medicine 2014; 6: 835-848.

457 Riviere J, Hauer J, Poirot L, Brochet J, Souque P, Mollier K *et al*. Variable correction of Artemis deficiency by I-Sce1-meganuclease-assisted homologous recombination in murine hematopoietic stem cells. Gene Ther 2014; 21: 529-532.

458 Rols MP. Electropermeabilization, a physical method for the delivery of therapeutic molecules into cells. Biochimica et biophysica acta 2006; 1758: 423-428.

459 Romero Z, Urbinati F, Geiger S, Cooper AR, Wherley J, Kaufman ML *et al*. beta-globin gene transfer to human bone marrow for sickle cell disease. J Clin Invest 2013.

460 Roncarolo MG, Battaglia M. Regulatory T-cell immunotherapy for tolerance to self antigens and alloantigens in humans. Nat Rev Immunol 2007; 7: 585-598.

461 Rosado IV, Langevin F, Crossan GP, Takata M, Patel KJ. Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. Nat Struct Mol Biol 2011; 18: 1432-1434.

462 Roselli EA, Mezzadra R, Frittoli MC, Maruggi G, Biral E, Mavilio F *et al*. Correction of beta-thalassemia major by gene transfer in haematopoietic progenitors of pediatric patients. EMBO molecular medicine 2010; 2: 315-328.

463 Rosenberg PS, Zeidler C, Bolyard AA, Alter BP, Bonilla MA, Boxer LA *et al*. Stable long-term risk of leukaemia in patients with severe congenital neutropenia maintained on G-CSF therapy. Br J Haematol 2010; 150: 196-199.

464 Rosselli F, Sanceau J, Gluckman E, Wietzerbin J, Moustacchi E. Abnormal lymphokine production: a novel feature of the genetic disease Fanconi anemia. II. In vitro and in vivo spontaneous overproduction of tumor necrosis factor alpha. Blood 1994; 83: 1216-1225.

465 Rothkamm K, Kruger I, Thompson LH, Lobrich M. Pathways of DNA double-strand break repair during the mammalian cell cycle. Mol Cell Biol 2003; 23: 5706-5715.

466 Rouet P, Smih F, Jasin M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol 1994; 14: 8096-8106.

467 Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A 1994; 91: 6064-6068.

468 Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome. Nat Rev Cancer 2011; 12: 51-58.

469 Sahin U, Kariko K, Tureci O. mRNA-based therapeutics--developing a new class of drugs. Nat Rev Drug Discov 2014; 13: 759-780.

470 San Filippo J, Sung P, Klein H. Mechanism of eukaryotic homologous recombination. Annu Rev Biochem 2008; 77: 229-257.

471 Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 2013; 32: 347-355.

472 Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. A transcription activator-like effector toolbox for genome engineering. Nat Protoc 2012; 7: 171-192.

473 Santiago Y, Chan E, Liu PQ, Orlando S, Zhang L, Urnov FD *et al*. Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. Proc Natl Acad Sci U S A 2008; 105: 5809-5814.

474 Sather BD, Romano Ibarra GS, Sommer K, Curinga G, Hale M, Khan IF *et al*. Efficient modification of CCR5 in primary human hematopoietic cells using a megaTAL nuclease and AAV donor template. Science translational medicine 2015; 7: 307ra156.

475 Sato K, Toda K, Ishiai M, Takata M, Kurumizaka H. DNA robustly stimulates FANCD2 monoubiquitylation in the complex with FANCI. Nucleic Acids Res 2012; 40: 4553-4561.

476 Sawyer SL, Tian L, Kahkonen M, Schwartzentruber J, Kircher M, University of Washington Centre for Mendelian G *et al*. Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. Cancer Discov 2015; 5: 135-142.

477 Saydaminova K, Ye X, Wang H, Richter M, Ho M, Chen H *et al*. Efficient genome editing in hematopoietic stem cells with helper-dependent Ad5/35 vectors expressing site-specific endonucleases under microRNA regulation. Mol Ther Methods Clin Dev 2015; 1: 14057.

478 Scharer OD. DNA interstrand crosslinks: natural and drug-induced DNA adducts that induce unique cellular responses. Chembiochem 2005; 6: 27-32.

479 Scheckenbach K, Morgan M, Filger-Brillinger J, Sandmann M, Strimling B, Scheurlen W *et al*. Treatment of the bone marrow failure in Fanconi anemia patients with danazol. Blood Cells Mol Dis 2012; 48: 128-131.

480 Schindler D, Hoehn H. Fanconi anemia mutation causes cellular susceptibility to ambient oxygen. American journal of human genetics 1988; 43: 429-435.

481 Schmid-Burgk JL, Schmidt T, Kaiser V, Honing K, Hornung V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. Nat Biotechnol 2013; 31: 76-81.

482 Schneider M, Chandler K, Tischkowitz M, Meyer S. Fanconi anaemia: genetics, molecular biology, and cancer - implications for clinical management in children and adults. Clin Genet 2015; 88: 13-24.

483 Schuler D, Kiss A, Fabian F. Chromosomal peculiarities and "in vitro" examinations in Fanconi's anaemia. Humangenetik 1969; 7: 314-322.

484 Schultz JC, Shahidi NT. Tumor necrosis factor-alpha overproduction in Fanconi's anemia. Am J Hematol 1993; 42: 196-201.

485 Schumann K, Lin S, Boyer E, Simeonov DR, Subramaniam M, Gate RE *et al*. Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. Proc Natl Acad Sci U S A 2015; 112: 10437-10442.

486 Sebastiano V, Maeder ML, Angstman JF, Haddad B, Khayter C, Yeo DT *et al*. In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. Stem Cells 2011; 29: 1717-1726.

487 Sejas DP, Rani R, Qiu Y, Zhang X, Fagerlie SR, Nakano H *et al*. Inflammatory reactive oxygen species-mediated hemopoietic suppression in Fancc-deficient mice. J Immunol 2007; 178: 5277-5287.

488 Sertic S, Pizzi S, Lazzaro F, Plevani P, Muzi-Falconi M. NER and DDR: classical music with new instruments. Cell Cycle 2012; 11: 668-674.

489 Sharan SK, Morimatsu M, Albrecht U, Lim DS, Regel E, Dinh C *et al*. Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking Brca2. Nature 1997; 386: 804-810.

490 Shigechi T, Tomida J, Sato K, Kobayashi M, Eykelenboom JK, Pessina F *et al*. ATR-ATRIP kinase complex triggers activation of the Fanconi anemia DNA repair pathway. Cancer Res 2012; 72: 1149-1156.

491 Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P *et al*. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. Curr Gene Ther 2011; 11: 11-27.

492 Sims AE, Spiteri E, Sims RJ, 3rd, Arita AG, Lach FP, Landers T *et al*. FANCI is a second monoubiquitinated member of the Fanconi anemia pathway. Nat Struct Mol Biol 2007; 14: 564-567.

493 Singh TR, Bakker ST, Agarwal S, Jansen M, Grassman E, Godthelp BC *et al*. Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. Blood 2009; 114: 174-180.

494 Singh TR, Saro D, Ali AM, Zheng XF, Du CH, Killen MW *et al*. MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. Molecular cell 2010; 37: 879-886.

495 Smeenk G, de Groot AJ, Romeijn RJ, van Buul PP, Zdzienicka MZ, Mullenders LH *et al*. Rad51C is essential for embryonic development and haploinsufficiency causes increased DNA damage sensitivity and genomic instability. Mutat Res 2010; 689: 50-58.

496 Smih F, Rouet P, Romanienko PJ, Jasin M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. Nucleic Acids Res 1995; 23: 5012-5019.

497 Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res 2000; 28: 3361-3369.

498 Smith JR, Maguire S, Davis LA, Alexander M, Yang F, Chandran S *et al*. Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. Stem Cells 2008; 26: 496-504.

499 Smogorzewska A, Matsuoka S, Vinciguerra P, McDonald ER, 3rd, Hurov KE, Luo J *et al*. Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell 2007; 129: 289-301.

500 Sobeck A, Stone S, Landais I, de Graaf B, Hoatlin ME. The Fanconi anemia protein FANCM is controlled by FANCD2 and the ATR/ATM pathways. J Biol Chem 2009; 284: 25560-25568.

501 Somyajit K, Subramanya S, Nagaraju G. Distinct roles of FANCO/RAD51C protein in DNA damage signaling and repair: implications for Fanconi anemia and breast cancer susceptibility. J Biol Chem 2012; 287: 3366-3380.

502 Song J, Yang D, Xu J, Zhu T, Chen YE, Zhang J. RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. Nat Commun 2016; 7: 10548.

503 Soulier J, Leblanc T, Larghero J, Dastot H, Shimamura A, Guardiola P *et al*. Detection of somatic mosaicism and classification of Fanconi anemia patients by analysis of the FA/BRCA pathway. Blood 2005; 105: 1329-1336.

504 Spangrude GJ, Heimfeld S, Weissman IL. Purification and characterization of mouse hematopoietic stem cells. Science 1988; 241: 58-62.

505 Spardy N, Duensing A, Hoskins EE, Wells SI, Duensing S. HPV-16 E7 reveals a link between DNA replication stress, fanconi anemia D2 protein, and alternative lengthening of telomere-associated promyelocytic leukemia bodies. Cancer Res 2008; 68: 9954-9963.

506 Stein S, Ott MG, Schultze-Strasser S, Jauch A, Burwinkel B, Kinner A *et al*. Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. Nat Med 2010; 16: 198-204.

507 Stein S, Scholz S, Schwable J, Sadat MA, Modlich U, Schultze-Strasser S *et al*. From bench to bedside: preclinical evaluation of a self-inactivating gammaretroviral vector for the gene therapy of X-linked chronic granulomatous disease. Hum Gene Ther Clin Dev 2013; 24: 86-98.

508 Stephan V, Wahn V, Le Deist F, Dirksen U, Broker B, Muller-Fleckenstein I *et al*. Atypical X-linked severe combined immunodeficiency due to possible spontaneous reversion of the genetic defect in T cells. N Engl J Med 1996; 335: 1563-1567.

509 Stoepker C, Hain K, Schuster B, Hilhorst-Hofstee Y, Rooimans MA, Steltenpool J *et al*. SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. Nat Genet 2011; 43: 138-141.

510 Strathdee CA, Gavish H, Shannon WR, Buchwald M. Cloning of cDNAs for Fanconi's anaemia by functional complementation. Nature 1992; 358: 434.

511 Sugasawa K. Regulation of damage recognition in mammalian global genomic nucleotide excision repair. Mutat Res 2010; 685: 29-37.

512 Sugawara N, Haber JE. Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. Mol Cell Biol 1992; 12: 563-575.

513 Sun N, Zhao H. Seamless correction of the sickle cell disease mutation of the HBB gene in human induced pluripotent stem cells using TALENs. Biotechnol Bioeng 2014; 111: 1048-1053.

514 Sun W, Ji W, Hall JM, Hu Q, Wang C, Beisel CL *et al*. Self-assembled DNA nanoclews for the efficient delivery of CRISPR-Cas9 for genome editing. Angew Chem Int Ed Engl 2015; 54: 12029-12033.

515 Sung P, Klein H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. Nat Rev Mol Cell Biol 2006; 7: 739-750.

516 Surralles J, Jackson SP, Jasin M, Kastan MB, West SC, Joenje H. Molecular cross-talk among chromosome fragility syndromes. Genes Dev 2004; 18: 1359-1370.

517 Svahn J, Dufour C. Fanconi anemia - learning from children. Pediatr Rep 2011; 3 Suppl 2: e8.

518 Symington LS. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. Microbiol Mol Biol Rev 2002; 66: 630-670, table of contents.

519 Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. Cell 1983; 33: 25-35.

520 Szymczak AL, Workman CJ, Wang Y, Vignali KM, Dilioglou S, Vanin EF *et al*. Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. Nat Biotechnol 2004; 22: 589-594.

521 Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H *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 1998; 17: 5497-5508.

522 Tan I, Ng CH, Lim L, Leung T. Phosphorylation of a novel myosin binding subunit of protein phosphatase 1 reveals a conserved mechanism in the regulation of actin cytoskeleton. J Biol Chem 2001; 276: 21209-21216.

523 Tan PL, Wagner JE, Auerbach AD, Defor TE, Slungaard A, Macmillan ML. Successful engraftment without radiation after fludarabine-based regimen in Fanconi anemia patients undergoing genotypically identical donor hematopoietic cell transplantation. Pediatr Blood Cancer 2006; 46: 630-636.

524 Tao Y, Jin C, Li X, Qi S, Chu L, Niu L *et al*. The structure of the FANCM-MHF complex reveals physical features for functional assembly. Nat Commun 2012; 3: 782.

525 Tavtigian SV, Simard J, Rommens J, Couch F, Shattuck-Eidens D, Neuhausen S *et al*. The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. Nat Genet 1996; 12: 333-337.

526 Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G *et al*. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med 2014; 370: 901-910.

527 Thoms KM, Kuschal C, Emmert S. Lessons learned from DNA repair defective syndromes. Exp Dermatol 2007; 16: 532-544.

528 Tieleman DP. Computer simulations of transport through membranes: passive diffusion, pores, channels and transporters. Clinical and experimental pharmacology & physiology 2006; 33: 893-903.

529 Timmers C, Taniguchi T, Hejna J, Reifsteck C, Lucas L, Bruun D *et al*. Positional cloning of a novel Fanconi anemia gene, FANCD2. Molecular cell 2001; 7: 241-248.

530 Tischkowitz M, Dokal I. Fanconi anaemia and leukaemia - clinical and molecular aspects. Br J Haematol 2004; 126: 176-191.

531 Tischkowitz M, Xia B. PALB2/FANCN: recombining cancer and Fanconi anemia. Cancer Res 2010; 70: 7353-7359.

532 Tischkowitz M, Winqvist R. Using mouse models to investigate the biological and physiological consequences of defects in the Fanconi anaemia/breast cancer DNA repair signalling pathway. J Pathol 2011; 224: 301-305.

533 Tischkowitz MD, Hodgson SV. Fanconi anaemia. J Med Genet 2003; 40: 1-10.

534 Tolar J, Adair JE, Antoniou M, Bartholomae CC, Becker PS, Blazar BR *et al*. Stem cell gene therapy for fanconi anemia: report from the 1st international Fanconi anemia gene therapy working group meeting. Mol Ther 2011; 19: 1193-1198.

535 Tolar J, Becker PS, Clapp DW, Hanenberg H, de Heredia CD, Kiem HP *et al*. Gene therapy for Fanconi anemia: one step closer to the clinic. Hum Gene Ther 2012; 23: 141-144.

536 Touzot F, Hacein-Bey-Abina S, Fischer A, Cavazzana M. Gene therapy for inherited immunodeficiency. Expert Opin Biol Ther 2014; 14: 789-798.

537 Touzot F, Moshous D, Creidy R, Neven B, Frange P, Cros G *et al*. Faster T-cell development following gene therapy compared with haploidentical HSCT in the treatment of SCID-X1. Blood 2015; 125: 3563-3569.

538 Trujillo JP, Surralles J. Savior siblings and Fanconi anemia: analysis of success rates from the family's perspective. Genet Med 2015; 17: 935-938.

539 Tsai SQ, Joung JK. What's changed with genome editing? Cell Stem Cell 2014; 15: 3-4.

540 Ulrich HD. Regulating post-translational modifications of the eukaryotic replication clamp PCNA. DNA Repair (Amst) 2009; 8: 461-469.

541 Unno J, Itaya A, Taoka M, Sato K, Tomida J, Sakai W *et al*. FANCD2 binds CtIP and regulates DNA-end resection during DNA interstrand crosslink repair. Cell Rep 2014; 7: 1039-1047.

542 Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S *et al*. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 2005; 435: 646-651.

543 Uziel O, Reshef H, Ravid A, Fabian I, Halperin D, Ram R *et al*. Oxidative stress causes telomere damage in Fanconi anaemia cells - a possible predisposition for malignant transformation. Br J Haematol 2008; 142: 82-93.

544 Valeri A, Martinez S, Casado JA, Bueren JA. Fanconi anaemia: from a monogenic disease to sporadic cancer. Clin Transl Oncol 2011; 13: 215-221.

545 van de Vrugt HJ, Eaton L, Hanlon Newell A, Al-Dhalimy M, Liskay RM, Olson SB *et al*. Embryonic lethality after combined inactivation of Fancd2 and Mlh1 in mice. Cancer Res 2009; 69: 9431-9438.

546 van Rensburg R, Beyer I, Yao XY, Wang H, Denisenko O, Li ZY *et al*. Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells. Gene Ther 2012; 20: 201-214.

547 Vasanthakumar A, Arnovitz S, Marquez R, Lepore J, Rafidi G, Asom A *et al*. Brca1 deficiency causes bone marrow failure and spontaneous hematologic malignancies in mice. Blood 2016; 127: 310-313.

548 Vaz F, Hanenberg H, Schuster B, Barker K, Wiek C, Erven V *et al*. Mutation of the RAD51C gene in a Fanconi anemia-like disorder. Nat Genet 2010; 42: 406-409.

549 Velazquez I, Alter BP. Androgens and liver tumors: Fanconi's anemia and non-Fanconi's conditions. Am J Hematol 2004; 77: 257-267.

550 Verlinsky Y, Rechitsky S, Schoolcraft W, Strom C, Kuliev A. Preimplantation diagnosis for Fanconi anemia combined with HLA matching. JAMA 2001; 285: 3130-3133.

551 Vierstra J, Reik A, Chang KH, Stehling-Sun S, Zhou Y, Hinkley SJ *et al*. Functional footprinting of regulatory DNA. Nat Methods 2015; 12: 927-930.

552 Vinciguerra P, Godinho SA, Parmar K, Pellman D, D'Andrea AD. Cytokinesis failure occurs in Fanconi anemia pathway-deficient murine and human bone marrow hematopoietic cells. J Clin Invest 2010; 120: 3834-3842.

553 Virts EL, Jankowska A, Mackay C, Glaas MF, Wiek C, Kelich SL *et al*. AluY-mediated germline deletion, duplication and somatic stem cell reversion in UBE2T defines a new subtype of Fanconi anemia. Hum Mol Genet 2015.

554 Wada T, Candotti F. Somatic mosaicism in primary immune deficiencies. Curr Opin Allergy Clin Immunol 2008; 8: 510-514.

555 Waisfisz Q, Morgan NV, Savino M, de Winter JP, van Berkel CG, Hoatlin ME *et al*. Spontaneous functional correction of homozygous fanconi anaemia alleles reveals novel mechanistic basis for reverse mosaicism. Nat Genet 1999; 22: 379-383.

556 Walsh CE, Fu K, Brecher M. Retroviral-mediated gene transfer for Fanconi anemia group A patients - a clinical trial. Blood 2001; 98.

557 Wang AT, Kim T, Wagner JE, Conti BA, Lach FP, Huang AL *et al*. A Dominant Mutation in Human RAD51 Reveals Its Function in DNA Interstrand Crosslink Repair Independent of Homologous Recombination. Molecular cell 2015; 59: 478-490.

558 Wang AT, Smogorzewska A. SnapShot: Fanconi anemia and associated proteins. Cell 2015; 160: 354-354 e351.

559 Wang D, Gao G. State-of-the-art human gene therapy: part II. Gene therapy strategies and clinical applications. Discov Med 2014; 18: 151-161.

560 Wang J, Exline CM, DeClercq JJ, Llewellyn GN, Hayward SB, Li PW *et al*. Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. Nat Biotechnol 2015; 33: 1256-1263.

561 Wang LC, Gautier J. The Fanconi anemia pathway and ICL repair: implications for cancer therapy. Crit Rev Biochem Mol Biol 2010; 45: 424-439.

562 Wang SC, Lin SH, Su LK, Hung MC. Changes in BRCA2 expression during progression of the cell cycle. Biochem Biophys Res Commun 1997; 234: 247-251.

563 Wang X, Peterson CA, Zheng H, Nairn RS, Legerski RJ, Li L. Involvement of nucleotide excision repair in a recombination-independent and error-prone pathway of DNA interstrand cross-link repair. Mol Cell Biol 2001; 21: 713-720.

564 Wang X, Kennedy RD, Ray K, Stuckert P, Ellenberger T, D'Andrea AD. Chk1-mediated phosphorylation of FANCE is required for the Fanconi anemia/BRCA pathway. Mol Cell Biol 2007; 27: 3098-3108.

565 Watanabe N, Mii S, Asai N, Asai M, Niimi K, Ushida K *et al*. The REV7 subunit of DNA polymerase zeta is essential for primordial germ cell maintenance in the mouse. J Biol Chem 2013; 288: 10459-10471.

566 Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S. Assembly of designer TAL effectors by Golden Gate cloning. PLoS One 2011; 6: e19722.

567 Weterings E, Chen DJ. The endless tale of non-homologous end-joining. Cell Res 2008; 18: 114-124.

568 Whitney MA, Royle G, Low MJ, Kelly MA, Axthelm MK, Reifsteck C *et al*. Germ cell defects and hematopoietic hypersensitivity to gamma-interferon in mice with a targeted disruption of the Fanconi anemia C gene. Blood 1996; 88: 49-58.

569 Wolff JA, Budker V. The mechanism of naked DNA uptake and expression. Adv Genet 2005; 54: 3-20.

570 Wong JC, Buchwald M. Disease model: Fanconi anemia. Trends Mol Med 2002; 8: 139-142.

571 Wong JC, Alon N, McKerlie C, Huang JR, Meyn MS, Buchwald M. Targeted disruption of exons 1 to 6 of the Fanconi Anemia group A gene leads to growth retardation, strain-specific microphthalmia, meiotic defects and primordial germ cell hypoplasia. Hum Mol Genet 2003; 12: 2063-2076.

572 Woodard LE, Wilson MH. piggyBac-ing models and new therapeutic strategies. Trends Biotechnol 2015; 33: 525-533.

573 Wu Y, Brosh RM, Jr. FANCJ helicase operates in the Fanconi Anemia DNA repair pathway and the response to replicational stress. Curr Mol Med 2009; 9: 470-482.

574 Wurtele H, Little KC, Chartrand P. Illegitimate DNA integration in mammalian cells. Gene Ther 2003; 10: 1791-1799.

575 Wyman C, Kanaar R. DNA double-strand break repair: all's well that ends well. Annu Rev Genet 2006; 40: 363-383.

576 Xia B, Sheng Q, Nakanishi K, Ohashi A, Wu J, Christ N *et al*. Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. Molecular cell 2006; 22: 719-729.

577 Xia B, Dorsman JC, Ameziane N, de Vries Y, Rooimans MA, Sheng Q *et al*. Fanconi anemia is associated with a defect in the BRCA2 partner PALB2. Nat Genet 2007; 39: 159-161.

578 Xie F, Ye L, Chang JC, Beyer AI, Wang J, Muench MO *et al*. Seamless gene correction of beta-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. Genome research 2014; 24: 1526-1533.

579 Xu P, Tong Y, Liu XZ, Wang TT, Cheng L, Wang BY *et al*. Both TALENs and CRISPR/Cas9 directly target the HBB IVS2-654 (C > T) mutation in beta-thalassemia-derived iPSCs. Sci Rep 2015; 5: 12065.

580 Yagasaki H, Adachi D, Oda T, Garcia-Higuera I, Tetteh N, D'Andrea AD *et al*. A cytoplasmic serine protein kinase binds and may regulate the Fanconi anemia protein FANCA. Blood 2001; 98: 3650-3657.

581 Yamamoto K, Ishiai M, Matsushita N, Arakawa H, Lamerdin JE, Buerstedde JM *et al*. Fanconi anemia FANCG protein in mitigating radiation- and enzyme-induced DNA double-strand breaks by homologous recombination in vertebrate cells. Mol Cell Biol 2003; 23: 5421-5430.

582 Yang Y, Kuang Y, Montes De Oca R, Hays T, Moreau L, Lu N *et al*. Targeted disruption of the murine Fanconi anemia gene, Fancg/Xrcc9. Blood 2001; 98: 3435-3440.

583 Yannaki E, Karponi G, Zervou F, Constantinou V, Bouinta A, Tachynopoulou V *et al*. Hematopoietic stem cell mobilization for gene therapy: superior mobilization by the combination of granulocyte-colony stimulating factor plus plerixafor in patients with beta-thalassemia major. Hum Gene Ther 2013; 24: 852-860.

584 Yu C, Liu Y, Ma T, Liu K, Xu S, Zhang Y *et al*. Small molecules enhance CRISPR genome editing in pluripotent stem cells. Cell Stem Cell 2015; 16: 142-147.

585 Yu J, Russell JE. Structural and functional analysis of an mRNP complex that mediates the high stability of human beta-globin mRNA. Mol Cell Biol 2001; 21: 5879-5888.

586 Yu X, Chen J. DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains. Mol Cell Biol 2004; 24: 9478-9486.

587 Zecca M, Strocchio L, Pagliara D, Comoli P, Bertaina A, Giorgiani G *et al*. HLA-haploidentical T cell-depleted allogeneic hematopoietic stem cell transplantation in children with Fanconi anemia. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation 2014; 20: 571-576.

588 Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P *et al*. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 2015; 163: 759-771.

589 Zhang F, Ma J, Wu J, Ye L, Cai H, Xia B *et al*. PALB2 links BRCA1 and BRCA2 in the DNA-damage response. Curr Biol 2009; 19: 524-529.

590 Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 2011; 29: 149-153.

591 Zhang H, Kozono DE, O'Connor KW, Vidal-Cardenas S, Rousseau A, Hamilton A *et al*. TGF-beta Inhibition Rescues Hematopoietic Stem Cell Defects and Bone Marrow Failure in Fanconi Anemia. Cell Stem Cell 2016; 18: 668-681.

592 Zhang J, Zhao D, Wang H, Lin CJ, Fei P. FANCD2 monoubiquitination provides a link between the HHR6 and FA-BRCA pathways. Cell Cycle 2008; 7: 407-413.

593 Zhang J, Walter JC. Mechanism and regulation of incisions during DNA interstrand cross-link repair. DNA Repair (Amst) 2014; 19: 135-142.

594 Zhang L, Hernan R, Brizzard B. Multiple tandem epitope tagging for enhanced detection of protein expressed in mammalian cells. Mol Biotechnol 2001; 19: 313-321.

595 Zhang X, Sejas DP, Qiu Y, Williams DA, Pang Q. Inflammatory ROS promote and cooperate with the Fanconi anemia mutation for hematopoietic senescence. J Cell Sci 2007; 120: 1572-1583.

596 Zhang X, Shang X, Guo F, Murphy K, Kirby M, Kelly P *et al*. Defective homing is associated with altered Cdc42 activity in cells from patients with Fanconi anemia group A. Blood 2008; 112: 1683-1686.

597 Zierhut HA, Tryon R, Sanborn EM. Genetic counseling for Fanconi anemia: crosslinking disciplines. J Genet Couns 2014; 23: 910-921.

598 Zou J, Sweeney CL, Chou BK, Choi U, Pan J, Wang H *et al*. Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting. Blood 2011; 117: 5561-5572.

599 Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH *et al*. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat Biotechnol 2015; 33: 73-80.

600 Zychlinski D, Schambach A, Modlich U, Maetzig T, Meyer J, Grassman E *et al*. Physiological promoters reduce the genotoxic risk of integrating gene vectors. Mol Ther 2008; 16: 718-725.

XIII. APPENDIX I:

SUPPLEMENTARY TABLES AND FIGURES

1. SUPPLEMENTARY TABLES FROM INTRODUCTION

1.1. FA genes described up to now

FA Gene	Prevalence	Chromosomal location	Protein Size (kDa)	Function
FANCA [152, 314]	64%	16q24.3	163	Member of the core complex. Phosphorylated by ATR kinase. Interacts with BRCA1.
FANCB [343]	2%	Xp22.2	95	Member of the core complex.
FANCC [510]	12%	9q22.3	63	Member of the core complex. Breast and ovarian cancer susceptibility gene.
FANCD1/BRCA2 [226, 525]	2%	13q12-13	380	Involved in HRR: recruits and supports RAD51 filament formation onto DNA. Breast and ovarian cancer susceptibility gene.
FANCD2 [529]	4%	3p25.3	155,162	Member of the FA ID2 complex. Monoubiquitinated by FANCL (core complex). Recruitment of FAN1 and FANCP. Phosphorylated by ATR and by ATM after DNA damage. Essential for downstream nucleolytic incisions and translesion synthesis repair events.
FANCE [115]	1%	6p21.3	60	Member of the core complex. Directly binds FANCD2. Contributes to core complex integrity, by promoting nuclear accumulation of FANCC protein and FANCA.
FANCF [116]	2%	11p15	42	Member of the core complex. Required for the assembly of the FA core complex: adaptor protein stabilizing A/G and E/C interaction.
FANCG/XRCC9 [114]	8%	9p13	68	Member of the core complex. Interacts with XRCC3 and FANCD1.
FANCI [129, 243, 293, 492, 499]	1%	15q26.1	140, 149	Member of the FA ID2 complex. Monoubiquitinated by FANCL (core complex) in a FANCD2 dependent manner. It is also phosphorylated after DNA damage. Essential for downstream nucleolytic incisions and translesion synthesis repair events.

FANCJ/ BRIP1/ BACH1 [294]	2%	17q23.1	140	Involved in HRR: 5'-3' DNA helicase. Binds the BCRT domain of BRCA1. Breast and ovarian cancer susceptibility gene.
FANCL [342]	0.4%	2p16.1	43	Member of the FA core complex. E3 ubiquitin ligase of FANCD2 and I. Recruitment of FANCT.
FANCM [344]	0.1%	14q21.3	250	Member of the core complex. DNA translocase and unwinding activities. Critical for ATR activation during ICL repair. Participates in DNA repair in other contexts. Breast and ovarian cancer susceptibility gene.
FANCN/PALB2 [440, 452, 577, 589]	0.7%	16p12.1	140	Involved in HRR: essential for stability and localization of BRCA2. Breast and ovarian cancer susceptibility gene.
FANCO/RAD51C [548]	0.1%	17q22	43	Involved in HRR: interact with RAD51 paralogs and is reported to be important for Holliday junction resolution. Breast and ovarian cancer susceptibility gene.
FANCP/SLX4/ BTBD12 [270, 509]	0.5%	16p13.3	200	Holliday junction resolvase, Scaffold/regulator of XPF-ERCC1, MUS81-EME1 and SLX1 nucleases.
FANCQ/ERCC4/XPF [38]	0.1%	16p13.12	104	Associates with ERCC1 to form a FANCP/SLX4-dependent ICL unhooking nuclease. Participates in NER independently of FANCP.
FANCR/RAD51 [557]	NR	15q15.1	37	Involved in HRR. Recombinase activity. Interacts with FANCS and FANCD1.
FANCS/BRCA1 [476]	NR	17q21	208	Involved in HRR: interacts with D2/I complex and BRCA2-PALB2. Inhibition of NHEJ. Removal of CMG (CDC45-MCM-GINS) complex during ICL repair. Breast and ovarian cancer susceptibility gene.
FANCT/UBE2T [215, 553]	NR	1q32.1	22.5	FA core complex. Acts as a specific E2 ubiquitin-conjugating enzyme by associating with E3 ubiquitin-protein ligase FANCL and catalyzing monoubiquitination of FANCD2.
FANCV/REV7/MAD2L2 [35]	NR	1p36	24	Extension TLS polymerase Pol ζ subunit. Acts repairing one strand of the DSB that would be used as a template of the HRR pathway to repair the other strand. It is also

	involved in mitotic checkpoint regulation, and DNA repair pathway choice.

Table S1. Update table of the human FANC genes. The table shows the human FANC genes name and their gene symbol synonyms, their prevalence (updated from The Fanconi Anemia Mutation Database, the Rockefeller University) [8], their chromosomal location, their protein size and their main functions. NR: not reported.

1.2. Mouse models of FA

FA gene (mouse orthologs)	Targeted disruption	Strain	Viability	Sensitivity to ICL agents	Sensitivity to INF-γ	Sensitivity to IR	Growth retardation	Fertility	Congenital abnormalities	HSPCs and hematological abnormalities	Gonadal/ Germ cell defects	Cancer predis-position
Single mutants												
Fanca [97]	Exons 4-7 with *lacZ-neoR*	Mixed 129/Ola and C57BL/6	Viable	Yes	Yes	N/A	No	Reduced	None	No anemia. Mild thrombocytope-nia. Reduced number of megakaryocyte progenitors *in vitro*.	Yes	Yes
Fanca [571]	Exons 1-6 with *lacZ*	C57BL/6	Viable	Yes	N/A	N/A	Yes	Reduced	Yes	No	Yes	Increased
Fancb (***Fancb*** $^{-/y}$) [135, 256]	Disruption of the *Fancb* locus with ZFN	C57BL/6	Viable	Yes	N/A	N/A	N/A	Infertility	N/A	No anemia. Reduction in the HSCs and in the number of hematopoietic colonies *in vitro*. Low repopulating ability.	Yes, FANCB is critical at early stages of germ cell development.	N/A
Fancc [93]	Exon 8 with *neoR*	Mixed 129/Sv and C57BL/6	Viable	Yes	N/A	N/A	Yes	Reduced	Yes	No anemia. Low repopulating ability.	Yes	Increased
Fancc [568]	Exon 9 with *neoR*	Mixed 129/Sv and C57BL/6	Viable	Yes	Yes	N/A	No	Reduced	None	No anemia. Reduced number of hematopoietic colonies *in vitro*.	Yes	No
Fancd1 (***Brca2*** $^{\Delta 27/\Delta 27}$) [339, 386]	Exon 27	Mixed 129/Sv and C57BL/6	Viable, perinatal lethality	Yes	N/A	Yes	N/A	Reduced	None	No anemia. Reduced number of hematopoietic colonies *in vitro*. Low repopulating ability.	None	Yes
Fancd1 [489]	Exon11	129/Sv	Embryonic lethality	N/A	N/A	N/A	Yes, *in utero*	N/A	N/A	N/A	N/A	N/A

Fancd2 [223]	Exon 26 and 27 with *neoR*	Mixed 129/S4 and C57BL/6	Viable, perinatal lethality	Yes	N/A	Similar to *Fancc* mutants	Yes	Reduced	Yes	N/A	Yes	Yes
Fancd2 [421]	Intron 1 with *puroR*	C57BL/6	Viable	Yes	N/A	N/A	N/A	Reduced	None	No anemia. Reduced number of HSPCs and hematopoietic colonies *in vitro*. Low repopulating ability.	Yes	Increased
Fance [164]	Intron 8 with self-inactivating lentivirus carrying a tyrosinase minigene	FVB/N	Viable	N/A	N/A	N/A	N/A	Infertility	N/A	N/A	Yes, ovarian dysplasia, severe lack of follicle.	No
Fancf [22]	Exon 1 with *neoR*	Mixed 129/Ola and FVB	Viable	Yes	N/A	N/A	No	Infertility	None	N/A	Yes, compromised follicle development and spermatoge-nesis.	Yes
Fancg [582]	Exons 2-9 with *neoR*	Mixed 129/Sv and FVB	Viable	Yes	N/A	Yes	No	Reduced	None	No	Yes, abnormal gonadal histology.	No
Fancg [276]	Exons 1-4 with *hygR*	Mixed 129/Ola and FVB	Viable	Yes	N/A	No	No	Reduced	None	No anemia. Reduced number of hematopoietic colonies *in vitro* with MMC.	Yes	No
***Fancl* (*Pog*)** [3]	Exons 4-14 with the goat *β-globin*	129/Sv and C57BL/6	Embryonic lethality	N/A	N/A	N/A	No	Reduced	None	No	Yes	N/A
Fancm [21]	Exon 2 with loxP *neoR*, followed by deletion of *neoR*	Mixed FVB and C57BL/6	Viable perinatal lethality animals	Yes	N/A	N/A	No	N/A	None	N/A	Yes, reduced number of developing follicles in ovaries, and absence of sperm and spermatogo-nia.	Yes

Fancn [46, 449]	Intron 1 with *β-geo*	Mixed 129/Ola and C57BL/6N; mixed 129/FVB	Embryonic lethality	N/A	N/A	N/A	Yes, *in utero*	N/A	N/A	N/A	N/A	N/A
Fanco [284]	Exons 2-3 with *loxP neoR*, followed by deletion of *neoR*	Mixed 129/Sv and C57BL/6	Embryonic lethality	N/A	N/A	N/A	N/A	Hypomorphic mice infertile	No	N/A	Yes, in hypomorphic mice.	N/A
Fanco [495]	Intron 5 with a G>A splice-site mutation	C57BL/6	Embryonic lethality	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Fancp [83]	Intron 3 with *β-geo* and *loxP*-mediated deletion of exon3	C57BL/6N	Viable, increased perinatal lethality	Yes	N/A	No	Yes	Reduced	Yes	No anemia. Mild cytopenia.	Yes	N/A
Fancr/ Rad51 $^{M1-/-}$ [301]	Deletion of nucleotides 413 to 530 that codes for amino acids 73 to 112 with a DV*puro* vector	Mixed 129/Sv and C57BL/6J	Embryonic lethality	N/A	N/A	Yes	N/A	N/A	N/A	N/A	N/A	N/A
Fancs/ Brca1 [547]	Intron 21 to 22.Mx1-Cre–mediated *Brca1* deletion	C57BL/6J	Viable	Yes	N/A	N/A	N/A	N/A	N/A	Cytopenias and BMF. Reduced number of hematopoietic colonies *in vitro* with MMC. Low repopulating ability.	N/A	Yes
Fancv/Rev7 [565]	Insertion of PGK-*neoR* cassette between exon 1 and 3	Mixed 129/Sv and C57BL/6J	Partial embryonic lethality	N/A	N/A	N/A	Yes	Infertility	Yes	N/A	Yes, germ cell aplasia in the testes and ovaries.	N/A
Fancv/ Rev7 ^{C70R} [262]	Missense mutation in the N terminus of *Rev7/Mad2l2* (C70R).	Mixed C57BL/6J and C3H/HeJ	Partial embryonic lethality	Yes	N/A	N/A	Yes	Infertility	N/A	N/A	Yes, in proliferation of primordial germ cells.	N/A
Usp1 [268]	Exon 3 with *loxP neoR*, followed by deletion of *neoR*	Mixed 129/Sv and C57BL/6J	Viable, increased perinatal lethality	Yes	N/A	Yes	Yes	Male infertility	Yes	Low repopulating ability. Hypoplasia in BM.	Yes	N/A

Double mutants												
Fanca* and *Fancc [396]	From crossing *Fanca* -exon 37 was removed- X *Fancc* [93]	Both 129/Sv and C57BL/6	Viable	Yes	No	Mild	No	Reduced	None	No anemia. Reduced number of hematopoietic colonies *in vitro* with and without MMC.	Variable	N/A
Fancc* and *Fancg [435]	From crossing *Fancc* [93] X *Fancg* [582]	C57BL/6J	Viable	Yes	N/A	N/A	N/A	N/A	N/A	Anemia and BMF.	N/A	Yes (AML and MDS)
Fancc and *p53* [161]	From crossing *Fancc* [93] X *Trp53* [246]	C57BL/6J	Viable	Yes	No	N/A	No	Infertility	N/A	No anemia	N/A	Yes
Fancc* and *Sod1 [203]	From crossing *Fancc* [93] X *Sod1* [451]	Mixed 129/Sv and C57BL/6J	Viable	Yes	N/A	N/A	No	Infertility	None	Hypocellular BM. Reduced number of hematopoietic colonies *in vitro*. Anemia and leukopenia.	NA	N/A
Fancd2* and *Fancg [453]	From crossing *Fancd2* (Alan D'Andrea lab) X *Fancg* [582]	Mixed 129/Sv and C57BL/6J	Embryonic lethality	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Fancd2* and *Aldh2 [173]	Mice from [287]	Mixed 129/Sv and C57BL/6J	Viable	Yes	N/A	N/A	Yes	Infertility	N/A	Aplastic anemia	N/A	Yes, susceptibility to develop AML
Fancd2* and *Mlh1 [545]	From crossing *Fancd2*[223] X *Mlh1* [20]	Mixed 129/Sv and C57BL/6J or C57BL/6J	Embryonic lethality	Attenua-ted	N/A	N/A	Yes, *in utero*	Infertility	N/A	N/A	N/A	N/A
Fancd2* and *p53 [224]	From crossing *Fancd2*[223] X *Trp53* [128]	129/Sv	Viable	Do not develop cell cycle arrest	N/A	N/A	N/A	Infertility	N/A	N/A	Yes	Yes

Table S2. Mouse models for FA and their corresponding phenotype. The indicated characteristics include the targeted disruption, the mouse strain, the sensitivity to ICL agents and/or to INF-γ and/or to ionizing radiation (IR). Moreover if the mice present growth retardation, infertility, congenital abnormalities, HSPCs and hematological defects, gonadal or germ cell defects and cancer predisposition are also indicated. *LacZ*, β-galactosidase gene; *hygR*, hygromycin B resistance gene; *neoR*, neomycin resistance gene; *puroR*, puromycin resistance gene; *β-geo*, β-galactosidase-*neo* fusion gene; N/A, not analysed; AML: acute myeloid leukemia; MLS: myelodisplastic syndrome.

1.3. Clinical trials conducted by GT with HSCs

Disease	Vector	Country (Patients)	Conditioning regime	Target Cell	Trans-duced Cells	Efficacy and Adverse Effects
HAEMATOLOGICAL DISORDERS						
ADA-SCID	γ-RV [34, 373, 379]	USA (2)	None	Lymphocytes T cells	Lymphocytes T cells (from PB)	Improvement immunologic parameters. Difficult to attribute to GT or ER. No AE.
	γ-RV [4, 5, 57, 180, 508]	Italy/Israel (18) USA (16) UK (8)	Busulfan Melphalan or Busulfan	Lymphoid cells and natural killer (NK) cells	HSCs (from BM)	Transient or no clinical benefit: 15/18 and 10/16 off ER. 4/8 off ER. No AE.
	SIN-LV	UK (14) USA (16)	Busulfan	Lymphoid cells and NK cells	HSCs (from BM or PB)	Promising results, improvement T cell numbers suggesting the procedure' safety and efficacy similar to those observed in previous trials [6]. No AE.
SCID-X1	γ-RV [99, 155, 179, 197, 199, 200, 225]	UK (11) France (11) USA (3)	None	Lymphoid cells and NK cells	HSCs (from BM and PB)	Significant sustained PB T-reconstitution functional to cope with infectious agents that are otherwise lethal in SCID (even those that developed leukemia). No clinical benefits in the 5 older patients: loss of thymic function. AE: T-ALL in 5 patients, 1 death.
	SIN γ-RV [201, 372, 537]	UK, France, USA (11)	None	Lymphoid cells and NK cells	HSCs (from BM)	Low but functional reconstitution of T-cells. No AE.
	SIN-LV	USA	Busulfan	Lymphoid cells and NK cells	HSCs (from PB or BM)	Ongoing.
X-CGD	γ-RV [29, 30, 188, 193, 253, 254, 329, 405, 506]	USA (13) Germany (2) Switzerland (2) UK (4) Korea (2)	None or Busulfan Busulfan Busulfan Melphalan Busulfan and Fludarabine	Neutrophils	HSCs (from PB)	No clinical benefits. Transient clinical improvement with no persistent efficacy. AE: 2 cases of MDS (1 death) and 2 develop mutagenic clonal expansion.
	SIN γ-RV [507]	Germany (0)	Busulfan	Neutrophils	HSCs (from PB)	Ongoing.
	SIN-LV	UK, Switzerland, Germany, France (1) USA (1)	Busulfan (and Fludarabine)	Neutrophils	HSCs (from PB)	Ongoing.
WAS	γ-RV [47, 48]	Germany (10)	Busulfan	Multi-lineage reconstitution	HSCs (from BM)	Clinical benefit, no immunodeficiency and reduced thrombocytopenia, correction in 9/10 in 2 year follow-up. AE: 6 cases of T-ALL and 1 case of AML (2 T-ALL develop secondary AML) (two deaths).
	SIN LV [6, 66, 202]	UK, USA, France (13) Italy (8)	Busulfan and fludarabine	Multi-lineage reconstitution	HSCs (from BM or PB)	Clinical benefit, no immunodeficiency, and reduced thrombocytopenia. No AE.

SCD	SIN LV [459]	USA (3)	Busulfan	Erythroid cells	HSCs (from BM)	Ongoing. One patient show early clinical benefit.
		USA	Unknown	Erythroid cells	HSCs (from BM)	Ongoing.
β-Thal	SIN LV [67, 163, 349, 389, 462]	France (3)	Busulfan	Erythroid cells	HSCs (from PB)	Therapeutic benefit. One patient became transfusion independence. AE: Clonal dominance upon vector integration within HMGA2.
		France (4)	Busulfan	Erythroid cells	HSCs (from PB)	Ongoing. Two patients independent of transfusions.
		Italy	Busulfan	Erythroid cells	HSCs (from BM or PB)	Ongoing (First patient recently treated).
		USA (10)&(4)	Busulfan	Erythroid cells	HSCs (from BM or PB)	Ongoing. Transfusion independence for the majority.
FA	γ-RV [261, 311, 556] (FANCA and FANCC)	USA (4) & (2)	None	HSCs	HSCs (from BM and PB)	No clinical benefit. Transient increase in Hb and PLT counts and in BM cellularity and HSCs (depending on the study). Few cells infused or low transduction efficiencies. Few cells detected or no engraftment. No AE.
	SIN LV [189, 534, 535]	USA (2)	None	HSCs	HSCs (from PB)	Ongoing.
		Spain (2)(with France, UK, and Germany)	None*			Ongoing.
STORAGE DISORDERS						
X-ALD	SIN LV [62]	UK, France (4)	Busulfan and CTX	Macrophages and microglia	HSCs	Therapeutic benefit comparable to HSCT. No AE.
		USA, France				
MLD	SIN LV [32]	Italy (9)	Busulfan	Macrophages and microglia	HSCs	Milder clinical benefit beyond the predicted age of symptom onset. No AE.

Table S3. Updated *ex vivo* HSCs GT clinical trials. Update from the website ClinicalTrials.gov and from recent publications [41, 71, 156, 185, 283, 384, 432, 536]. ADA-SCID: Adenosine deaminase severe combined immunodeficiency; SCID-X1: X-linked SCID; X-CGD: X-linked chronic granulomatous disease; WAS: Wiskott-Aldrich syndrome; β-Thal: β-thalassemia; SCD: sickle cell disease; FA: Fanconi anemia; MLD: metachromatic leukodistrophy; ALD: adrenoleukodystrophy; NK: natural killer cells; γ-RV: gamma retroviral vectors; SIN γ-RV: self-inactivating gamma retroviral vectors; SIN-LV: self-inactivating lentiviral vectors; ER: enzyme replacement; T-ALL: T cell acute lymphoblastic leukemia; AML: acute myeloid leukemia; MDS: myelodysplastic syndrome; Hb: hemoglobin; PLT: platelets; AE: adverse effects; CTX: cyclophosphamide;*: if no engraftment is observed, the next group of patients may receive conditioning before cell infusion (fludarabine in combination with cyclophosphamide or ionizing radiation or even antibody-based regiments).

2. SUPPLEMENTARY TABLES FROM MATERIALS AND METHODS

2.1. Summary of the experiments performed in Lin$^-$ BM cells

Cocktail 1: IMDM with 20 % FBS. Factors: 100 ng/ml of mSCF, hIL11, hFlt3, human granulocyte colony stimulating factor (hGCSF), hTPO and 20 ng/ml of mouse fibroblast growth factor basic (mFGF2).
Cocktail 2: IMDM with 20 % FBS. Factors: 100 ng/ml of mSCF, hIL11, hFlt3, hTPO.
Cocktail 3: StemSpan™. Factors: 100 ng/ml of mSCF, hIL11, hFlt3, hTPO.
Cocktail 4: StemSpan™ & Glu 1%. Factors: 100 ng/ml of mSCF, hIL11, hFlt3, hTPO.

Experiment	Nucleofector (L, large or S, strips cuvettes)	Kit	Program	Cells/ condition	DNA /mRNA (µg)	Strain	Age of mice (w, weeks)	Culture conditions
1	DeviceI-L	Mouse Macrophages	X-01	$1x10^6$	2	FVB WT	8-10 w	Cocktail 1
2	DeviceI-L	Mouse Macrophages	U1, U8, U15, A13	$1.5x10^6$	2	FVB WT	8-10 w	Cocktail 1
3	4D-L	P3	EO-100, ED-113, ET-113	$2.3x10^6$	2-2.5	FVB WT	8-10 w	Cocktail 1
4	4D-L	P3	EO-100, ED-113, ET-113	$1.2x10^6$	2-2.5	FVB WT	8-10 w	Cocktail 1
5	4D-S	P3	ED-113, DT-113, DS-118, EH-100, EO-100, DS-109.	$0.2\text{-}1x10^6$	0.4	FVB WT	8-10 w	Cocktail 1
6	4D-S	P3	DS118	$0.5x10^6$	2/0.4-2	FVB WT	8 w	Cocktail 2
7	4D-L	P3	ED-113	$2x10^6$	2-2.5	FVB WT	8-10 w	Cocktail 2
8	4D-L	P3	ED-113	$1x10^6$	2.5-10	FVB WT	9-10 w	Cocktail 3
9-12; 21-23	4D-L	P3	ED-113	$1\text{-}1.3x10^6$	2.5-9	FVB FA	8-12 w	Cocktail 3
13	4D-L	P3	ED-113	$1\text{-}1.4x10^6$	2.5-9	FVB FA	8-10 w	Cocktail 3 + QVD
14-15;17	4D-L	P3	ED-113	$1\text{-}1.4x10^6$	2.5-9	FVB FA	8-10 w	Cocktail 3 + QVD
16	4D-L	P3	ED-113	$0.8x10^6$	/3-6-12	FVB FA	8-10 w	Cocktail 3 + QVD
18	4D-L	P3	ED-113	$1.6x10^6$	1.5-5	FVB FA	8-10 w	Cocktail 3 + QVD
24	4D-L	P3	ED-113	$1.4x10^6$	/2.5-50	FVB FA	14 w	Cocktail 3
25-26	4D-L	P3	ED-113, DN-110, EA-105, ED-123, ED-113, DS123, EN113.	$1\text{-}1.3x10^6$	2.5-9	FVB FA	8-12 w	Cocktail 3
27-29	4D-L	P3	ED-113	$1.1\text{-}1.4x10^6$	1.5-9	C57BL/6 WT	8-10 w	Cocktail 3 + QVD
30; 32 ; 34	4D-L	P3	ED-113	$1.4x10^6$	1.5-5.5-9	C57BL/6 WT	8-10 w	Cocktail 3
35-38	4D-L	P3	ED-113	$1.4x10^6$	4-9	C57BL/6 WT	8-10 w	Cocktail 4 + QVD
39	4D-L	P3	ED-113	$1.4x10^6$	4-9	C57BL/6 WT	10-14 w	Cocktail 4 + QVD
40	4D-L	P3	DK-100	$1x10^6$	2.5/4	C57BL/6 WT	10 w	Cocktail 4 + QVD

Table S4. Summary of the different experiments of gene targeting performed in Lin$^-$ BM cells. Indicating the nucleofector machine, the kit of nucleofection, the program, the cells nucleofected per condition, the delivery method (DNA or mRNA), the strain of mice use as a source of Lin$^-$ cells, the age of mice and the culture conditions.

3. SUPPLEMENTARY FIGURES AND TABLES FROM RESULTS

3.1. Nucleotide alignment by BLAST of the sequencing of the 3' integration junction in clone #76# of FA-A MEFs

- **Right mHA of therapeutic cassette**
- ***Mbs85* locus (m*AAVS1)***

Query: Sequence derived from SequencingAAVS1-3_3'_F.

Subject: Plasmid Sequence with the 3' junction.

Length: 1336 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
1120 bits (606)	0.0	609/610 (99%)	1/610 (0%)	Plus/Plus

```
Query  2     GAGG-CTACATGATCTGCCGGTGTACCCTGGGAATGAGTTCCCAGTGTTGTCTGGGACAT  60
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  727   GAGGCCTACATGATCTGCCGGTGTACCCTGGGAATGAGTTCCCAGTGTTGTCTGGGACAT  786

Query  61    TTTCCTTTACATACAGCAGGGCACTACTTCCCTGCGGCCCCTCTCAGGGTCTCAGCATCT  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  787   TTTCCTTTACATACAGCAGGGCACTACTTCCCTGCGGCCCCTCTCAGGGTCTCAGCATCT  846

Query  121   TTAAACAGTGCAGGAGAGCAGTTCAGACTTCTAGCACCACCCAAGTCCCTGGCTTCTGAG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  847   TTAAACAGTGCAGGAGAGCAGTTCAGACTTCTAGCACCACCCAAGTCCCTGGCTTCTGAG  906

Query  181   CTTTCTCTTCTGACTGCATCCCTCTCCTCTGCATCTGCTGGAGGCACTAGAGCACCTCCT  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  907   CTTTCTCTTCTGACTGCATCCCTCTCCTCTGCATCTGCTGGAGGCACTAGAGCACCTCCT  966

Query  241   CTGACAGGACCTGTAACCTCAGGCCTGGCCTGTCAGGCAGCCTTCCAGGTGGGTCTCTCC  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  967   CTGACAGGACCTGTAACCTCAGGCCTGGCCTGTCAGGCAGCCTTCCAGGTGGGTCTCTCC  1026

Query  301   TCCCACCGCTGCGATACCTTGGTGAACTCACTCTTCAGGGCTTTGCAGGAGGCTTGGATG  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1027  TCCCACCGCTGCGATACCTTGGTGAACTCACTCTTCAGGGCTTTGCAGGAGGCTTGGATG  1086
```

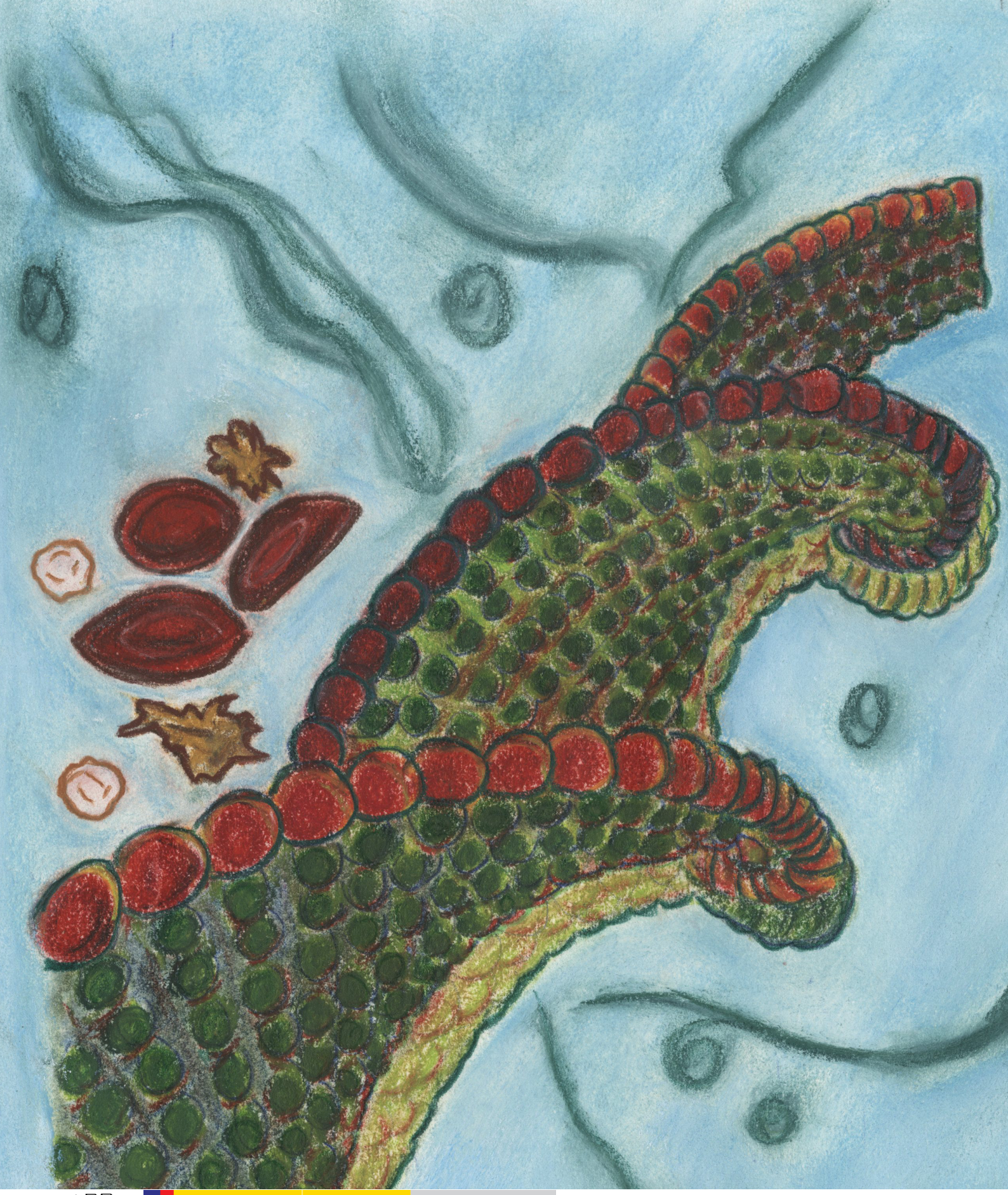
```
Query  361   ATTGGGAAAGGAGATGCAAAGAGAGCAGTTGGTTTACAGGGCCAGTACGCAGAGGGAGCT  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1087  ATTGGGAAAGGAGATGCAAAGAGAGCAGTTGGTTTACAGGGCCAGTACGCAGAGGGAGCT  1146

Query  421   TGGGTCTAGAGTGGGGTGAGTGTGCTGTCACCACCTTGTACTGCCCTGGCTTTAGCAGTG  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1147  TGGGTCTAGAGTGGGGTGAGTGTGCTGTCACCACCTTGTACTGCCCTGGCTTTAGCAGTG  1206

Query  481   AAAGTTCCAAGAGCTAATTTCAGTCCCAGGAAAACAAGGACATTTGGTCACTCTGTCTGG  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1207  AAAGTTCCAAGAGCTAATTTCAGTCCCAGGAAAACAAGGACATTTGGTCACTCTGTCTGG  1266

Query  541   TGACTTCTGGTGTCTCCTGATTTCTATACCTGGAGCTGTGTTTTCAGTTTTAAACCCTCC  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1267  TGACTTCTGGTGTCTCCTGATTTCTATACCTGGAGCTGTGTTTTCAGTTTTAAACCCTCC  1326

Query  601   TTGGGGCATC  610
             ||||||||||
Sbjct  1327  TTGGGGCATC  1336
```

Figure S1. Nucleotide alignment by BLAST of the 3' PCR integration junction sequence. Sequencing of the 3' integration junction obtained by PCR representing part of the therapeutic cassette with designed right mHA (green) and part of the right location of the *Mbs85* locus (blue). Representative blast using as a query the sequence derived from SequencingAAVS1-3_3'_F primer and as a subject the *Mbs85* sequence corresponding to the 3' integration junction.

3.2. Summary of the transplants performed with gene-edited WT Lin$^-$ BM (CD45.2$^+$) cells in P3B recipient mice (CD45.1$^+$)

Experiments	Conditions	Number of nucleofected transplanted cells	Number of transplanted mice that died
Evaluation of the *in vivo* repopulating ability of gene-edited cells in P3B (CD45.1$^+$) mice		**Exp1 (Lin$^-$ 29)**	
	MOCK (ø)	110,000	0/2
	T (2,5 µg)+D (4 µg)	148,000	0/1
	D (4 µg)	20,000	1/1
		Exp2 (Lin- 30)	
	MOCK (ø)	366,000	0/3
		122,500	
		48,000	
	T (0,75 µg)+D (4 µg)	122,500	0/1
	T (2,5 µg)+D (4 µg)	366,000	0/1
	D (4 µg)	48,000	0/1
		Exp3 (Lin$^-$ 32)	
	T (0,75 µg)+D (4 µg)	122,000	0/1
	D (4 µg)	142,000	1/1
		Exp4 (Lin$^-$ 34)	
	T (0,75 µg)+D (4 µg)	626,000	0/1
	D (4 µg)	586,000	0/1
Evaluation of longer periods of expansion in the engraftment potential of gene-edited cells in P3B (CD45.1$^+$) mice		**Exp5 (Lin$^-$ 35)**	
	MOCK (ø)	161,568	1/1
		335,000	0/1
	T (2,5 µg)+D (4 µg)	161,568	1/1
	D (4 µg)	335,000	1/1
		Exp6 (Lin$^-$ 37)	
	T (2,5 µg)+D (4 µg)	190,008	1/1
	D (4 µg)	354,983	0/1
		Exp7 (Lin$^-$ 38)	
	T (2,5 µg)+D (4 µg)	685,000	1/1
	D (4 µg)	825,000	1/1
		685,000	0/1
		500,000	1/1
Evaluation of the effect of molecules that could enhance genome editing and the engraftment potential of gene-edited cells in P3B (CD45.1$^+$) mice		**Exp8 (Lin$^-$ 39)**	
	MOCK (ø)	85,000	0/3
		250,000	
		1,000,000	
	T (2,5 µg)+D (4 µg)	85,000	0/1
	T (2,5 µg)+D (4 µg)+ L755507	65,000	1/1
	T (2,5 µg)+D (4 µg)+ BrefeldinA	250,000	1/1

Table S5. Analysis of the *in vivo* repopulating ability of gene-edited WT Lin$^-$ BM (CD45.2$^+$) cells in P3B mice (CD45.1$^+$). Table recapitulates eight different experiments of transplant performed with gene-edited or mock C57BL/6 Lin$^-$ cells transplanted in P3B mice, indicating the number of cells and the number of transplanted mice that died within the first week of the transplant due to graft failure.



UAM UNIVERSIDAD AUTÓNOMA DE MADRID

GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

IIS FJD
INSTITUTO DE INVESTIGACIÓN SANITARIA FUNDACIÓN JIMÉNEZ DÍAZ

ciberer isciii