UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS
Departamento de Bioquímica y Biología Molecular

TESIS DOCTORAL

Terapia génica dirigida en un nuevo "sitio seguro" en células progenitoras hematopoyéticas humanas para su aplicación en anemia de Fanconi

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

Fátima Rodríguez Fornés

Directores

Juan A. Bueren Roncero
Guillermo Güenechea Amurrio

Madrid, 2016

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA CELULAR

TERAPIA GÉNICA DIRIGIDA EN UN NUEVO "SITIO SEGURO" EN CÉLULAS PROGENITORAS HEMATOPOYÉTICAS HUMANAS PARA SU APLICACIÓN EN ANEMIA DE FANCONI

FÁTIMA RODRÍGUEZ FORNÉS

TESIS DOCTORAL

MADRID 2015

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA CELULAR

Targeted Gene Therapy into a New "Safe Harbor" Site in Human Hematopoietic Progenitor Cells for Applications in Fanconi Anemia

Memoria presentada por FÁTIMA RODRÍGUEZ FORNÉS licenciada en Biología, para optar al grado de doctor por la Universidad Complutense de Madrid, con mención europea.

Directores de Tesis

Dr. Juan A Bueren Roncero

Dr. Guillermo Güenechea Amurrio

Madrid, 2015

CENTRO DE INVESTIGACIONES ENERGÉTICAS, MEDIO AMBIENTALES Y TECNOLÓGICAS (**CIEMAT**)
CENTRO DE INVESTIGACIONES BIOMÉDICAS EN RED DE ENFERMEDADES RARAS (**CIBERER**)
INSTITUTO DE INVESTIGACIÓN SANITARIA DE LA FUNDACIÓN JIMÉNEZ DÍAZ (**IIS – FJD**)

Guillermo Güenechea Amurrio, científico titular de la División de Terapias Innovadoras del Sistema Hematopoyético del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), y **Juan Antonio Bueren Roncero**, jefe de la misma División, certifican que la memoria adjunta titulada "*Terapia génica dirigida en un nuevo "sitio seguro" en células progenitoras hematopoyéticas humanas para su aplicación en anemia de Fanconi"* ha sido realizada por Fátima Rodríguez Fornés bajo la dirección conjunta de los que suscriben, y cumple las condiciones exigidas para optar al título de Doctor con mención europea por la Universidad Complutense de Madrid.

Dr. Juan Antonio Bueren Roncero

Dr. Guillermo Güenechea Amurrio

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

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 (Proyectos SAF2009-07164 y SAF2012-39834).
- 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.

Fátima Rodríguez Fornés ha disfrutado de una beca de Formación de Personal Investigador (FPI) del Ministerio de Economía y Competitividad (Ref BES-2010-037339) y de un contrato del IIS-FJD.

Acute myeloid leukemia (AML)

Adeno-associated viruses (AAVs)

Adenosine Deaminase-Severe Combined Immunodeficiency (ADA-SCID)

Base Pairs (bp)

Bone Marrow (BM)

Bone Marrow Failure (BMF)

Clustered Regularly Interspaced Short Palindromic Repeats Cas9 system (CRISPR/Cas9)

Colony Forming Units (CFU)

Dyepoxibutane (DEB)

Double Strand Break (DSB)

Human Elongation Factor-1 alpha (EF1α)

Embryonic stem cells (ESCs)

Eukaryotic elongation factor 1 alpha (EF1α)

Enhanced Green Fluorescence Protein (EGFP)

Erythropoietin (EPO)

Fluorescence-Activated Cell Sorting (FACS)

Fanconi Anemia (FA)

FMS-Like Tyrosine kinase 3 ligand (Flt3)

Forward (Fw)

Gene Therapy (GT)

Genomic DNA (gDNA)

Genomic Safe Harbors (GSHs)

Graft-Versus-Host Disease (GVHD)

Granulocyte Colony-Stimulating Factor (G-CSF)

Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF)

Chronic Granulomatous Disease (CGD)

Green Matrix (GM)

HEK-293H

Hematopoietic Stem Cell (HSC)

Hematopoietic Stem Cells Transplantation (HSCT)

Hematopoietic Stem Progenitor Cells (HSPCs)

Homing Endonuclease (HE)

Homology Arms (HA)

Homologous Recombination (HR)

Induced Pluripotent Stem Cells (iPSCs)

Insertions-Deletions (INDELS)

Integrase-Deficient Lentivirus Vectors (IDLVs)

Interferon gamma (IFNγ)

Interleukin-1β (IL-1β)

Interleukin-3 (IL-3)

Interleukin-6 (IL-6)

Interstrand Cross-Links (ICLs)

Lentivirus (LV)

Long Terminal Repeats (LTRs)

Mean Fluorescent Intensity (MFI)

Meganuclease (MN)

Mitomicyn-c (MMC)

Multipotent Progenitors (MPP)

Myelodysplastic Syndrome (MDS)

Non Homologous End Joining (NHEJ)

NOD/SCID/IL2rγ^{null} (NSG)

Nucleotide Excision Repair (NER)

Penicillin/Streptomycin (P/S)

Peripheral blood (PB)

Protospacer-Adjacent Motif (PAM)

Relative Humidity (RH)

Reverse (Rv)

Right-Homology Arm (R-HA)

Safe Harbor (SH)

Safe Harbor 6 (SH6)

Safe Harbor 6 Meganuclease (SH6-MN)

Safe Harbor 6 TALEN (SH6-TALEN)

Stem Cell Factor (SCF)

SCID Repopulating Cells (SRCs)

Self Inactivating (SIN)

Severe Combined Immunodeficiency (SCID)

Squamous Cell Carcinoma (SCC)

Stem Cell Factor (SCF)

Transcription Activator-Like Effector Nucleases (TALENs)

Thrombopoietin (TPO)

Therapeutic Matrix (TM)

Three Prime Repair Exonuclease 2 (TREX2)

Translesion Synthesis (TLS)

Tumor Necrosis Factor alpha (TNFα)

Umbilical Cord Blood (UCB)

Wild Type (wt)

Wiskott-Aldrich Syndrome (WAS)

X-linked Severe Combined Immunodeficiency (SCID-X1)

Zinc-Finger Nucleases (ZFN)

INDEX

RESUMEN

I. RESUMEN

La anemia de Fanconi es una enfermedad hereditaria de baja prevalencia, descrita por primera vez por el pediatra Guido Fanconi en 1927. Esta enfermedad se produce como consecuencia de mutaciones en cualquiera de los 19 genes de Fanconi descritos hasta la actualidad, y que participan en la ruta de Fanconi/BRCA. Esta ruta se encarga de la reparación de enlaces intercatenarios del ADN y de coordinar los distintos mecanismos de reparación de las dobles roturas en el ADN. La anemia de Fanconi está caracterizada por generar inestabilidad genómica, lo que da lugar a anomalías esqueléticas y predisposición al cáncer, si bien la principal causa de muerte de pacientes pediátricos es el fallo de médula ósea.

Uno de los tratamientos alternativos al trasplante alogénico de progenitores hematopoyéticos de pacientes con anemia de Fanconi se basa en la reinfusión de células madre hematopoyéticas autólogas, tras su corrección con vectores lentivirales. Para limitar al máximo los riesgos de este tipo de terapias se están desarrollando nuevas tecnologías de edición génica basadas en la inserción dirigida de los genes terapéuticos. Esta nueva aproximación se fundamenta en la generación de dobles roturas en regiones específicas del genoma, cuya reparación por recombinación homóloga facilitaría la entrada de los genes terapéuticos aportados por ADNs donadores externos con homología por dicha región.

En este trabajo se ha desarrollado una aproximación de edición génica en un nuevo "sitio seguro" del genoma denominado SH6. Para ello se ha trabajado con la línea celular HEK-293H, así como también con progenitores hematopoyéticos humanos purificados en base a la expresión del marcador CD34. Para su desarrollo se han utilizado nucleasas de edición, tales como meganucleasas y TALEN, en combinación con matrices donadoras portadoras del gen marcador *EGFP* (GM) o del gen terapéutico *FANCA* (TM). En todos los casos los genes marcadores y terapéuticos estaban regulados por el promotor EF1α, y flanqueados por dos brazos de homología para el sitio SH6. Estos plásmidos han servido como molde para realizar la terapia génica de edición en el sitio seguro SH6.

Con objeto de optimizar la edición génica con las nucleasas seleccionadas se comenzó a trabajar en la línea celular HEK-293H que fue transfectada en primer lugar con la matriz portadora del marcador *EGFP* (GM) y dos meganucleasas con especificidad por el sitio SH6 (SH6v2 y SH6v5). Se cuantificó el porcentaje de células que expresaban el transgén en cultivo a lo largo del tiempo, observándose en aproximadamente el 4% de las células se mantenía la expresión de la proteína EGFP. Se observó también que ambas meganucleasas eran capaces de generar integraciones específicas del donador GM en el sitio SH6 (hasta un 9,89%). Tras comprobar el establecimiento de pseudoclones positivos para esta integración específica, se verificó que la expresión del transgén se mantenía estable en el tiempo. Se obtuvieron, asimismo, resultados similares cuando se comparó la eficacia de la meganucleasa SH6v5 respecto a la TALEN SH6.

En experimentos posteriores comprobamos la generación de células editadas utilizando la matriz terapéutica TM portadora del gen *FANCA*, junto con la meganucleasa SH6v5 o la TALEN

SH6. Estos experimentos demostraron una mayor eficacia de la meganucleasa SH6v5, si bien esta eficacia fue siempre inferior a la obtenida con la matriz portadora del gen *EGFP*, probablemente por su mayor tamaño.

Una vez optimizada la edición génica en células HEK-293H, aplicamos nuestra aproximación experimental a progenitores hematopoyéticos humanos $CD34^+$. En primer lugar se realizaron estudios en los que se determinó la dosis máxima de ADN con la que nucleofectar estas células, y se verificó su capacidad clonogénica después del proceso de nucleofección y separación de las células fluorescentes por FACS. Asimismo, se cuantificó la proporción de progenitores células primitivas después del tratamiento, y se seleccionó el mejor momento para la nucleofección con nucleasas en forma de ADN o ARNm. Puesto que la preselección de células $EGFP^+$ tras la nucleofección no permitió generar colonias hematopoyéticas con inserciones en el sitio SH6, se realizaron nuevos experimentos en los que no se realizó ninguna preselección de la población nucleofectada. En estas condiciones, la nucleofección de las células $CD34^+$ con 5 µg de ADN total de la SH6-TALEN y 4 µg del donador GM generó un 2,1% de colonias editadas; si bien también se observó una significativa citotoxicidad. La utilización del nuevo nucleofector 4D, o el uso de la TALEN SH6 como ARNm, modificada o sin modificar en su extremo 3´UTR no mejoró los resultados obtenidos. Finalmente, la nucleofección en el equipo AMAXA I de las células $CD34^+$ con 10 µg de ADN total de la TALEN SH6 y 4 µg de la matriz GM permitió obtener un 3,13% de colonias hematopoyéticas en las que se había insertado el gen marcador en el sitio seguro SH6.

En conjunto, nuestros resultados sugieren que el sitio SH6 constituye un sitio seguro de interés para llevar a cabo terapia génica dirigida en progenitores y células madre hematopoyéticas humanas. Estas observaciones tienen un impacto particular para la terapia génica dirigida de enfermedades tales como la anemia de Fanconi, en donde la ventaja proliferativa de un número reducido de células madre corregidas podría tener un impacto clínico significativo.

SUMMARY

II. SUMMARY

Fanconi anemia is an inherited disease of low prevalence that was first described by the pediatrician Guido Fanconi in 1927. This disease occurs as a result of mutations in any of the 19 Fanconi genes so far described, and involved in the Fanconi/BRCA pathway. This pathway is responsible for the repair of the DNA interstrand cross-links and coordinates different repair mechanisms of DNA double strand breaks. Fanconi anemia is characterized by genome instability, and results in skeletal abnormalities and predisposition to cancer, although the main cause of death in pediatric patients is the bone marrow failure.

One of the alternative treatments to allogeneic hematopoietic stem cell transplantation in Fanconi anemia patients is based on the reinfusion of autologous hematopoietic stem cells, after correction with lentiviral vectors. In order to minimize the risks of such therapies, new gene editing technologies based on the targeted integration of the therapeutic genes are under development. These new approaches are based on the generation of double strand breaks in specific regions of the genome that facilitates the homologous recombination of therapeutic genes provided by external DNA homology donors.

In this work we have developed an approach for gene editing in a new "safe harbor" of the genome called SH6. To achieve this goal, we have worked with the HEK-293H cell line, as well as with purified human hematopoietic progenitors that expressed the CD34 marker. For this development, specific nucleases such as meganucleases and TALEN were used in combination with donor matrixes that carried the *EGFP* marker gene (GM) or the therapeutic *FANCA* gene (TM). In all cases the marker and therapeutic genes were regulated by the EF1α promoter, and flanked by two homology arms for SH6 site. These plasmids have served as templates for targeted gene therapy on the safe harbor SH6.

In order to optimize the gene editing with the selected nucleases, HEK-293H cell line was transfected first with the matrix carrying the *EGFP* marker (GM) and two meganucleases specific for the SH6 site (SH6v2 and SH6v5). The percentage of cells expressing the transgene in culture over time was quantified and approximately 4% of cells maintained the expression of the EGFP protein. It was also observed that both meganucleases were able to generate specific integration of the GM donor in the SH6 site (up to 9.89%). After confirming the establishment of positive pseudoclones for this specific integration, it was verified that the transgene expression was stable over time. Similar results were also obtained when the efficacy of the SH6v5 meganuclease was compared to the SH6 TALEN.

In subsequent experiments we evaluated the generation of edited cells using the therapeutic matrix TM carrying the *FANCA* gene along with the SH6v5 meganuclease or the TALEN SH6. These experiments demonstrated the higher efficacy of the SH6v5 meganuclease, although this efficiency was always lower than that obtained with the matrix carrying the *EGFP* gene, probably because of their larger size.

Once we optimized gene editing in HEK-293H cells, we applied our experimental approach to human $CD34^{+}$ hematopoietic progenitors. First, the highest doses of DNA to nucleofect these

cells were determined. Also their clonogenic capacity after nucleofection process and separation of the fluorescent cells by FACS was verified. Thereafter, the proportion of primitive progenitor cells after treatment was determined, and the optimal time for $CD34^+$ cells nucleofection with the nucleases, used either as DNA or mRNA, was selected. Since the screening of $EGFP^+$ cells after nucleofection did not generate hematopoietic colonies with inserts in SH6 site, new experiments in which no pre-selection of the nucleofected population were performed. Under these conditions, the nucleofection of $CD34^+$ cells with 5 µg of total DNA of TALEN SH6 and 4 µg of GM donor generated 2.1% of edited colonies; although a significant cytotoxicity was also observed in these experiments. The use of the new nucleofector 4D, or the use of SH6 TALEN as mRNA, modified or unmodified at its end 3'UTR, did not improve the results. Finally, the nucleofection with the AMAXA nucleofector I of $CD34^+$ cells with 10 µg of total DNA of TALEN SH6 and 4 µg of GM matrix yielded 3.13% of hematopoietic colonies in which the marker gene was inserted in the safe harbor SH6.

Taken together, our results suggest that the SH6 site is a relevant safe harbor locus to be considered in targeted gene therapy in human hematopoietic progenitors and stem cells. Our observations may have a particular impact for targeted gene therapy of diseases such as Fanconi anemia, where the proliferative advantage of a small number of corrected stem cells could have a significant clinical impact.

INTRODUCTION

1. FANCONI ANEMIA

1.1 General characteristics of Fanconi anemia disease

Fanconi anemia (FA) is a rare hereditary genetic disease that was described for the first time by the Swiss pediatrician Guido Fanconi in 1927. Fanconi anemia is suffered by 1-3 per 1,000,000 habitants and thus is classified as a rare disease (Antonio Casado, Callen et al. 2007). FA is characterized by genome instability disorder that gives rise to skeletal abnormalities and cancer predisposition (D'Andrea and Grompe 1997), inducing acute myeloid leukemia (AML), myelodysplastic syndrome (MDS) and squamous cell carcinoma (SCC) in head, neck and oral cavities (Howlett, Taniguchi et al. 2005). However, the bone marrow failure (BMF) is the main characteristic of the disease and the main cause of mortality of the patients.

This disease is caused by the mutation in any of the 19 genes that nowadays are known to be implicated in the FA/BRCA pathway. From these 19 genes, 15 are classified as bona fide FA genes following a stringent criteria in which at least two patients with BMF and a positive chromosome fragility test should be identified (*FANCA, B, C, D1, D2, E, F, G, I, J, L, N, P, Q and T*);Other 3 genes are known as FA-like genes, because they cause a chromosome fragility syndrome and FA related malformations but without BMF (*FANCO, FANCR and FANCS*). There are serious doubts if *FANCM* should be considered as a FA gene (Bogliolo and Surralles 2015) due to the fact that only a patient is known to carry this mutation, but in combination with biallelic FANCA mutation. Even more, both *FANCM* and *FANCA* genes are necessary for the full correction of the phenotype (Singh, Bakker et al. 2009) and the *FANCM* loss-of-function variants are more common than originally predicted because two individuals of Finnish descent with loss of *FANCM* function variants are healthy and exhibit normal hematology(Lim, Wurtz et al. 2014). All of the genes show an autosomal recessive inheritance, but not *FANCB*, which is linked to the X chromosome (Adam, Azoulay et al. 2003, Meetei, Levitus et al. 2004). All these genes are involved in the FA/BRCA pathway, required for the repair of the DNA interstrand cross-links (ICLs), and also to be a potential regulator of DNA repair pathways of double strand breaks (DSBs), including homologous recombination (HR) (Kee and D'Andrea 2010, Bogliolo and Surralles 2015).

The most common FA-gene mutated in Spain is *FANCA*, which is present in 80% of the diagnosed patients. Among them, there is an important proportion of gypsy patients who share an ancestral mutation in exon 4 (Callen, Casado et al. 2005, Antonio Casado, Callen et al. 2007).

1.2 The Fanconi anemia/BRCA pathway

DNA is duplicated by DNA polymerases during the S phase of the cell cycle at multiple replication forks. Therefore, lesions in the DNA induced as ICLs could impair the replication

process at these forks. The FA/BCRA pathway is essential for the repair of the ICLs. All the genes nowadays implicated in the FA-pathways and their complementation groups are represented in **Table 1**. All known FA proteins are classified in three groups as described below **Figure 1**:

- ## Core complex

Eight FA proteins (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, and FANCM) form a nuclear complex (FANCore), characterized by having an ubiquitin E3 ligase that is activated by blocked DNA replication forks (Kennedy and D'Andrea 2005, Bogliolo and Surralles 2015). The heterodimer FANCM-FAAP24 is responsible for the recognition of the DNA damage (Ciccia, Ling et al. 2007) and its association with chromatin is mediated by proteins MHF1 and MHF2 (histone fold protein 1 and 2) (Singh, Saro et al. 2010) and by the DNA damage signaling mediated by ATR (ataxia telangiectasia and Rad3-related) (Collis, Ciccia et al. 2008, Schwab, Blackford et al. 2010). Even if FANCM is not functional, the mismatch repair complex MutS is able to recognize DNA ICLs, due to its redundant activity, activating the FA-pathway (Huang, Kennedy et al. 2011), and although the consideration of FANCM as a bona fide FA-gene is in question, this protein is important in the FA-pathway playing a critical role and working in collaboration with other FA core members because the complex FANCM/FAAP24 works as a sensor that recognize DNA damage and recruits the FA core complex that will induce the monoubiquination of FANCD2/FANCI (Kee and D'Andrea 2010).

Once FANCM/FAAP24 recognizes the DNA lesion, the core complex is recruited through the binding to FANCF. This core complex is constituted by two subcomplexes, FANCA/FANCG (in which FANCA is critical for the half-life of FANCG, that interacts with FANCF (Garcia-Higuera, Kuang et al. 2000) and with FANCB, and the FANCE/FANCF/FANCG/FANCC subcomplex. FANCB interacts with FANCL, which contains the PHD E3 ubiquitin ligase domain that will monoubiquinitate its two substrates: FANCD2 and FANDI, the “ID-complex” (Kennedy and D'Andrea 2005, Kee and D'Andrea 2010).

- ## ID-complex

The ID-complex is formed by the proteins FAND2 and FANCI, which are monoubiquinated by the E3 ubiquitin ligase domain of FANCL in cooperation with FANCT (UBE2T E2) (Smogorzewska, Matsuoka et al. 2007, Virts, Jankowska et al. 2015). This monoubiquinated ID-complex moves onto chromatin and interacts with other DNA repair proteins, such as BRCA1 or RAD51 and with other downstream FA-proteins (D1, J, N, O, P and Q) (Kennedy and D'Andrea 2005, Kee and D'Andrea 2012, Kottemann and Smogorzewska 2013).

- ## Downstream proteins

Once ID-complex has been monoubiquinated and has moved onto chromatin, it recruits several proteins involved in the ICLs repair, such as FAAN1 (*FA-associated nuclease* 1) and FANCP/SLX4 (FANCP) (D'Andrea and Grompe 1997, Crossan and Patel 2012). These proteins will recruit the downstream FA-proteins and other proteins involved in the DNA repair. FANCP/SLX4 works as a scaffold that will interact with three proteins (XPF, Mus81 and SLX1)

that make incisions in both sides of the covalently linked nucleotides of the ICL. Thereafter, the *Translesion synthesis (TLS)* takes place; a damage tolerance process in which TLS polymerases extend the master strand and restores it so that could be used as a template for the HR (Chang and Cimprich 2009). The *Homologous Recombination* process involves the downstream FA proteins FANCD1/BRCA2, FANCJ/BRIP1, FANCN/PALB2 and FANCO/RAD51C, and maybe FANCP/SLX4, regulating MUS81 and SLX1 (Kottemann and Smogorzewska 2013). This process will repair the DSB using the homologous DNA strand repaired by TLS as a template to restore the DNA with a high fidelity. The last step is the repair of the formed adducts by *Nucleotide excision repair (NER)*. Finally, DNA polymerases fill the nicks that remain in the DNA.

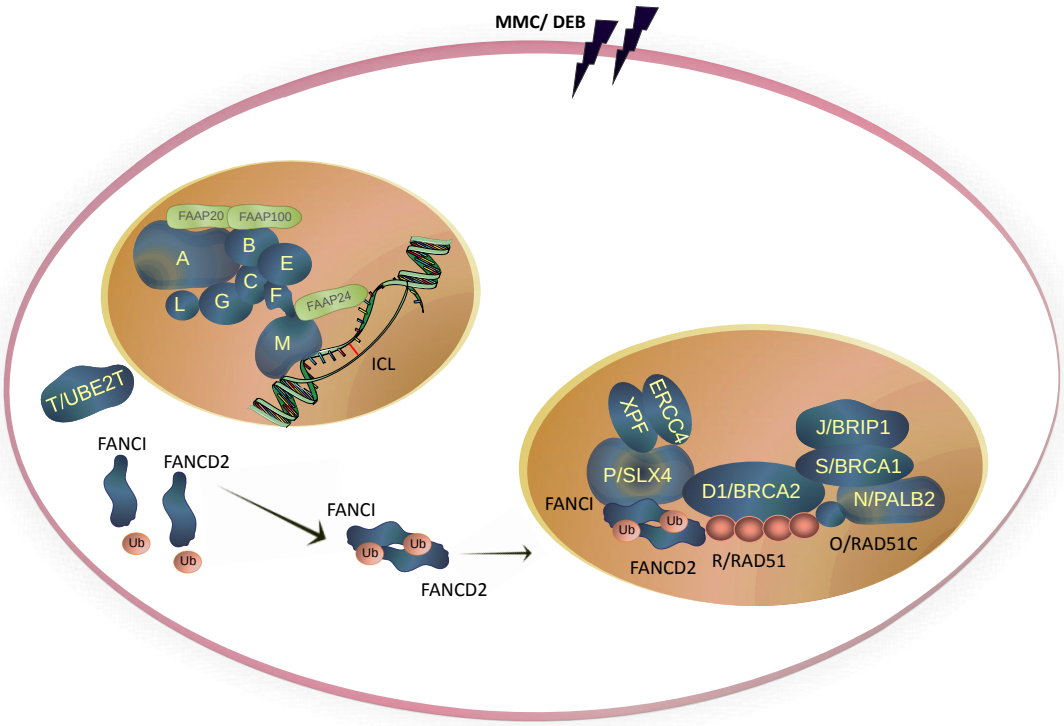

Figure 1: FA/BRCA pathway, constituted by all the proteins organized in three groups: The Core complex, the ID-complex and the downstream proteins, activated after ICLs lesions induced in the DNA by, for example, mitomycine (MMC) or dyepoxybutane (DEB) drugs.

Gene	%	Locus	Ub-D2/I	Function	Reference
FANCA	64	16q24.3	Yes	Scaffold for FA core complex	(Fanconi anaemia/Breast cancer 1996, Lo Ten Foe, Rooimans et al. 1996)
FANCB	2	Xp22.31	Yes	Scaffold for FA core complex	(Meetei, Levitus et al. 2004)
FANCC	12	9q22.3	Yes	Scaffold for FA core complex	(Strathdee, Gavish et al. 1992)
FANCE	1	6p21-22	Yes	Scaffold for FA core complex and adaptor for FAND2	(de Winter, Leveille et al. 2000)
FANCF	2	11p15	Yes	Scaffold for FA core complex	(de Winter, Rooimans et al. 2000)
FANCG XRCC9	8	9p13	Yes	Scaffold for FA core complex	(de Winter, Waisfisz et al. 1998)
FANCL PHF9	0.4	2p16.1	Yes	FA core complex. Recruits FANCT/UBE2T. Monoubiquinates FANCD2 and I	(Meetei, de Winter et al. 2003)
FANCM HEF	0.1	14q21.3	Yes	Scaffold for FA core complex. DNA translocase. Critical for ATR activation during ICL repair. Predisposes to breast and ovarian cancer.	(Meetei, Medhurst et al. 2005)
FANCT UBE2T	0.1	1q32.1	Yes	FA core complex. Provides the E2 conjugase activity to ubiquinate D2/I	(Hira, Yoshida et al. 2015, Rickman, Lach et al. 2015, Virts, Jankowska et al. 2015)
FANCD2	2	3q25.3	Yes	Monoubiquinated by FANCL. Heterodimer with FANDI. Recruits FAN1 and FANCP	(Timmers, Taniguchi et al. 2001)
FANCI	1	15q25-26	Yes	Monoubiquinated by FANCL. Heterodimer with FANCD2	(Dorsman, Levitus et al. 2007)
FANCO RAD51C	0.1	17q22	No	HR. Interacts with RAD51. Predisposes to breast and ovarian cancer.	(Vaz, Hanenberg et al. 2010)
FANCP SLX4	0.5	16p13.3	No	Holiday junction resolvase, scaffold/regulator of XPF-ERCC1, MUS81-EME1 and SLX1.	(Kim, Lach et al. 2011)
FANCQ ERCC4 XPF	0.1	16p13.12	No	Associates with ERCC1 to form FANCP/SLX4-dependent ICL unhooking nuclease. NER.	(Bogliolo, Schuster et al. 2013)

Gene	%	Locus	Ub-D2/I	Function	Reference
FANCD1 BRCA2	2	13q12-13	No	HR. Recruits RAD51. Predisposes to breast and ovarian cancer.	(Howlett, Taniguchi et al. 2002)
FANCR RAD51	0.1	15q15.1	No	Critical in HR, recombinase activity. Interacts with FANCs and D1.	(Wang, Kim et al. 2015)
FANCS BRCA1	0.1	17q21	No	HR. Interacts with D2/I complex and BRCA2-PALB2. Inhibition of NHEJ. Predisposes to breast and ovarian cancer.	(Sawyer, Tian et al. 2015)
FANCJ BRIP1	2	17q22-24	No	Interacts with BRCA1, HR: 5´-3´ DNA helicase, ATPase. Predisposes to breast and ovarian cancer.	(Levitus, Waisfisz et al. 2005)
FANCN PALB2	0.7	16p12.1	No	Mediates interaction BRCA1-BRCA2 in HR. Predisposes to breast and ovarian cancer.	(Reid, Schindler et al. 2007)

Table 1: Fanconi anemia genes are represented in this table with their location in the genome, other denominations and main function and activities that they perform.

1.3 Phenotype of Fanconi anemia cells

Fanconi anemia is caused by mutations in any of the FA genes showed in **Table 1**. However, there is controversy about the origin of its phenotypic manifestation. These manifestations could be due to the defect of the repair of the ICLs-DNA damage that gives rise to an accumulation of DNA damage and genomic instability. Additionally this phenotype could be originated by the imbalance in the oxygen metabolism and to an increase in the oxidative stress. Most probably the combination of these two circumstances may be the origin of the FA phenotype (Pallardo, Lloret et al. 2010, Pagano, Talamanca et al. 2013).

Since FA cells have a characteristic defect in the repair of ICLs, the use of **crosslinking agents** as mitomicyn c (MMC) or dyepoxibutane (DEB) is particularly toxic for FA cells due to the accumulation of chromosomal aberrations. This property is thus used for the diagnosis of the disease (Schuler, Kiss et al. 1969, Auerbach, Rogatko et al. 1989, Castella, Pujol et al. 2011).

FA cells present defective *ex-vivo* growth properties and low clonogenic capacity due to the imbalance in the oxygen metabolism (Joenje, Arwert et al. 1981, Saito, Hammond et al. 1993), originating **oxidative stress**. Some FA-proteins (as for example: FANCA, C, G, D2 and J) are suggested to be implicated in both the FA-pathway and in the redox detoxification metabolism

(Zhang, Sejas et al. 2007). Recently a defective mitochondrial function that produces a defective cellular energy metabolism has been studied in FA-A cells (Ravera, Vaccaro et al. 2013). FANCD2 also participates in oxygen damage response and FANCJ can inhibit the oxygen damage sensor hemoxigenase 1 (Pagano, Talamanca et al. 2012).

The accumulation of blocked DNA forks induced by ICL drugs like MMC or DEB, induce a **cell cycle arrest in G2/M** (Kennedy and D'Andrea 2005). The **apoptosis is increased** in FA cells due to the production of pro-apoptotic cytokines as interleukin 6 (IL-6), tumor necrosis factor alpha (TNFα), interleukin 1β (IL-1β) or interferon gamma (IFNγ), and also to the hypersensitive of FA cells to these molecules. The homeostasis of hematopoietic stem cells (HSCs) and their progeny is severely affected as well as their proliferation properties and their differentiation in the bone marrow niche thus leading to the BMF characteristic of FA patients(Du, Erden et al. 2014). The **homing and adhesion** of these cells are affected due to the CDC42 and CXCR5 activity (Skinner, O'Neill et al. 2008, Zhang, Shang et al. 2008).

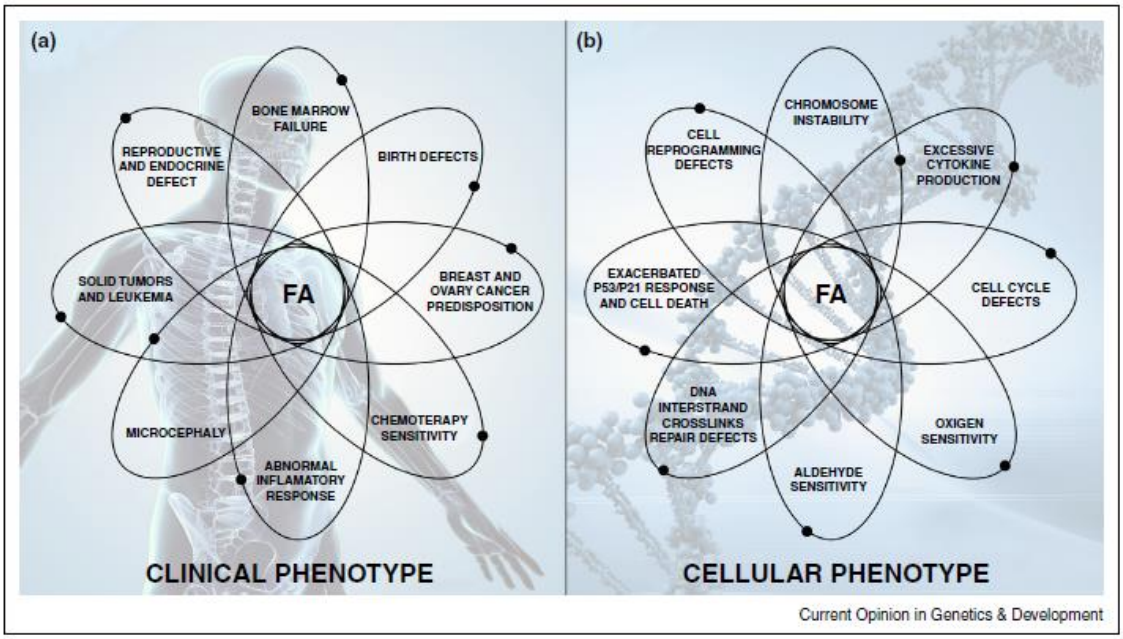

Figure 2: The panel (a). shows all the clinical characteristics of FA patients, all of them with incomplete penetrance. In (b) the main cellular consequences of the defective DNA repair mechanisms in FA cells are shown. Picture from Bogliolo *et al.* 2015 (Bogliolo and Surralles 2015).

1.4 Diagnosis, complementation groups and conventional treatments

Although many FA patients present a common symptomatology, the clinical manifestations of the disease are highly variable. The correlation between the mutation and phenotype of each patient has been tried to be evaluated. However, only a few incidences have been stablished, because this association is modulated by the context, the age and other genetic alterations potentially affecting the phenotype of these patients (Neveling, Endt et al. 2009).

- Diagnosis

The first differential test that was performed for FA patients is based on the cell exposure to ICL agents, such as MMC and DEB. After the exposure the presence of chromosomal breaks is studied (Auerbach, Rogatko et al. 1989). Also the G2/M arrest in MMC-treated cells is used, particularly in cultured fibroblasts.

In some patients an spontaneous recovery of the BMF can take place. The cause of this phenomenon, known as **somatic mosaicism,** is the generation of additional spontaneous mutations that can repair one of the affected alleles in hematopoietic precursors, now able to produce a functional FA protein. If this reversion of the mutation takes place in a HSC, it will be able to restore the hematological parameters by clonal expansion thanks to its proliferative advantage, replacing the FA-defective cells in the bone marrow. This phenomenon is considered a "Natural gene therapy" (Gross, Hanenberg et al. 2002, Antonio Casado, Callen et al. 2007).

- Complementation groups

Once the FA diagnosis is confirmed, it is convenient the identification of the mutatet FA gene that accounts for the disease. To this aim, peripheral blood T-cells can be transduced with different viral vectors (γ-retroviral or lentiviral) encoding a different FA-gene. The FA gene that corrects the hypersensitivity to MMC indicates the FA complementation group. In case that none of these most common FA genes correct the phenotype, Western Blot (WB) of FANCD2 can be performed to elucidate if the mutated gene is upstream or downstream in the FA pathway (Antonio Casado, Callen et al. 2007).

More recently, the development of massive sequencing approaches are facilitating the identification of the FA genes, and also of their specific mutation, according for the disease of most FA patients (Chandrasekharappa, Lach et al. 2013).

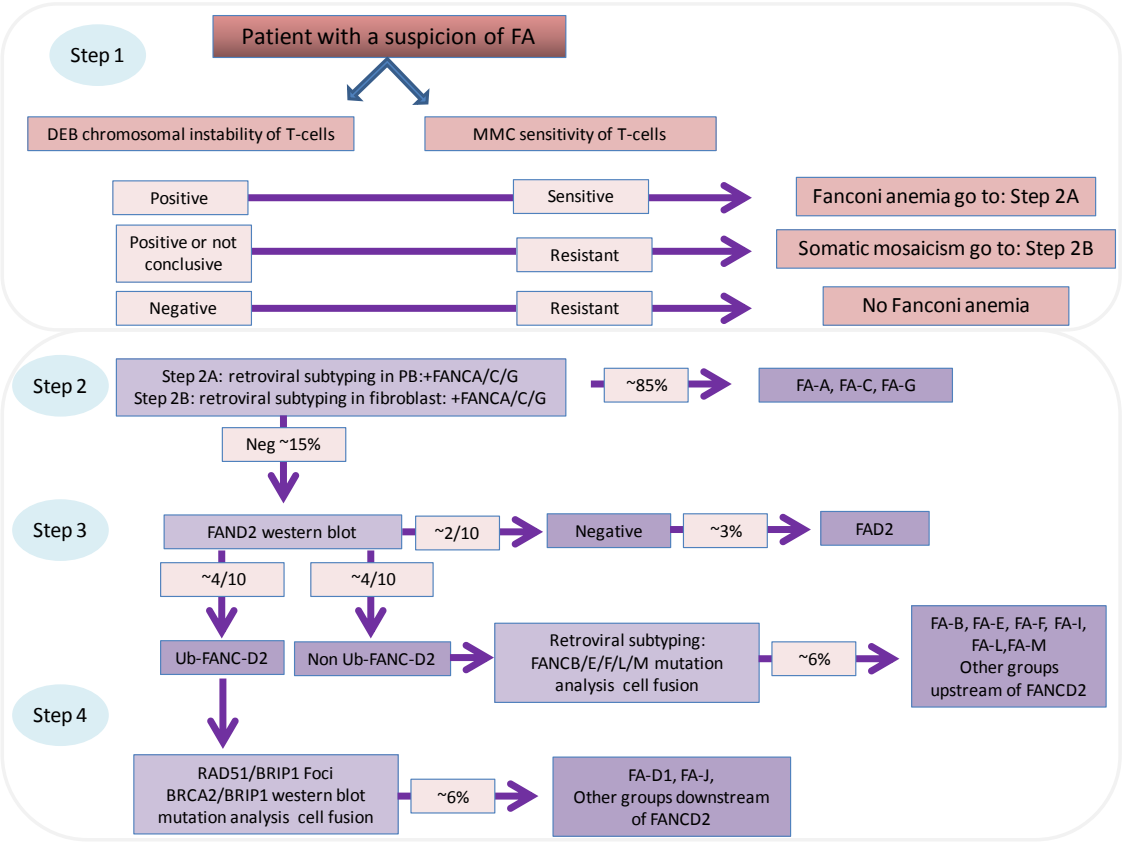

Figure 3: This scheme represents a four step procedure to diagnose FA-patients. In Step 1 the sensitivity of T-cells to MMC and DEB is measured. In the Step 2 the patients are subtyped using γ-RV carrying most of the FA-genes. Once the most common genes are elucidated to not account for the disease, in step 3 is evaluated if the mutated gene is up or downstream of the FA-pathway. Finally, in step 4, the involvement of the less common mutated FA-genes is studied. Picture modified from Casado *et al.* (Antonio Casado, Callen et al. 2007).

- Conventional treatment

The conventional treatments of FA patients include the use of pharmacological therapies. For example, **growth factors** as interleukin-3 (IL-3), erythropoietin (EPO), granulocyte-macrophage colony-stimulating factor (rhGM-CSF) or granulocyte colony-stimulating factor (rhG-CSF) aiming to increase the levels of the hematopoietic cells. Nevertheless, because the use of these growth factors could give rise to cell transformation, patients are not frequently treated with these molecules (Dufour and Svahn 2008).

Androgen therapy is frequently used in the initiating stages of the BMF. Most patients (75%) have a positive response to this treatment. Oxymetholone has been frequently used, although its long-term use can generate significant side effects. Danazol is a synthetic androgen derivative that has been also used in the last years with lower toxicity and similar efficacy (Velazquez and Alter 2004, Scheckenbach, Morgan et al. 2012).

When the BMF is severe, blood transfusions are necessary although they constitute a limiting factor for the success of **hematopoietic stem cell transplantation (HSCT)**. HSCT is the only therapy proven to be functional to restore the hematopoiesis of FA patients. The success of this therapy is particularly high in patients with less than 10 years old without expansions of aberrant clones (MacMillan and Wagner 2010, Svahn and Dufour 2011, Peffault de Latour, Porcher et al. 2013). Ideal donors are HLA-matched sibling donors, where the risk of graft-versus-host disease (GVHD)is lower (MacMillan and Wagner 2010, Peffault de Latour, Porcher et al. 2013). In the last years, the use of fludarabine in the conditioning regimen improved the outcome of HLA-matched related and also unrelated donors (Tan, Wagner et al. 2006). Also, the use of T-cell depletion has decreased the GVHD incidence (Deeg, Socie et al. 1996, Guardiola, Socie et al. 2004, Bonfim, de Medeiros et al. 2007, MacMillan, DeFor et al. 2015). The infusion of megadoses of HSC and the use of reduced intensity conditioning also improved the success in transplanted FA-patients (Zecca, Strocchio et al. 2014).

When FA-patients lack a HLA-compatible donor, the **selection of healthy HLA-identical pre-implantational embryos** constitutes a new experimental approach considered in the therapy of FA. In this case, healthy embryos that are HLA-identical to the FA patient are selected and implanted in the mother. HSCs from the umbilical cord blood of from the BM of these childs could be thus transplanted to the FA sibling. However, only a few numbers of patients have been treated by this very expensive and inefficient method (Bielorai, Hughes et al. 2004, Grewal, Kahn et al. 2004).

2. GENE THERAPY

Gene therapy (GT) is a novel and promising therapy based in the transfer of nucleic acids (DNA or RNA) to treat or prevent human diseases (Kaufmann, Buning et al. 2013). For this purpose integrative or not integrative viral vectors are used. This therapy can be performed either *in vivo* or *ex vivo*, but in the case of hematopoietic GT, the approaches that are more frequently used are *ex vivo* because cells obtained from the patient can be directly exposed to the vector and then re-infused after the correction. Nowadays this approach is preferred to *in vivo* therapy mainly to improve the transduction efficacy of the HSCs.

Gene therapy has shown clinical efficacy in a variety of monogenic diseases as the adenosine deaminase-severe combined immunodeficiency (ADA-SCID), Wiskott-Aldrich syndrome (WAS), chronic granulomatous disease (CGD), adrenoleukodystrophy, metachromatic leukodystrophy and hemoglobinopathies (**Table 2**).

Disease	Vector	Patients	Conditioning regime	Outcome	References
ADA	γ-RV	Italy (18), USA (14)	Busulfan	15/18 and 10/14 ERT	(Stephan, Wahn et al. 1996, Aiuti, Cattaneo et al. 2009, Gaspar, Cooray et al. 2011, Candotti, Shaw et al. 2012)
		UK (8)	Mephalan o Busulfan	4/8 ERT	
	SIN-LV	UK(10), USA (4)	Busulfan	Unpublished, on going	-
SCID-X1	γ-RV	UK (11), France (11), USA (3)	None	Significant T-reconstitution/T-ALL in 5 patients, 1 dead	(Hacein-Bey-Abina, Le Deist et al. 2002, Chinen, Davis et al. 2007, Hacein-Bey-Abina, Garrigue et al. 2008, Howe, Mansour et al. 2008, Hacein-Bey-Abina, Hauer et al. 2010, Gaspar, Cooray et al. 2011)
	SIN γ-RV	UK, France, USA (8)	None	Unpublished	(Hacein-Bey-Abina, Pai et al. 2014, Touzot, Moshous et al. 2015)
	LV	USA	Busulfan	On going	-
CGD	γ-RV	USA (5), and (3)	None or Busulfan	No benefits	-
		Germany/Switzerland (4)	Busulfan	Long-term correction 4/4	
		UK (4)	Melphalan	Transient correction	
		Korea (2)	Busulfan +Fludarabine	Short term correction	
	SIN γ-RV	UK, Switzerland, Germany	Busulfan	On-going	(Malech, Maples et al. 1997, Goebel and Dinauer 2003, Ott, Schmidt et al. 2006, Bianchi, Hakkim et al. 2009, Kang, Choi et al. 2010, Stein, Ott et al. 2010, Bianchi, Niemiec et al. 2011, Kang, Bartholomae et al. 2011)
		France	Busulfan/Fludara-bine	On-going	
	LV	UK, Switzerland, Germany, France, USA	Busulfan	On going	-
SCD	SIN LV	USA	Busulfan	Ongoing	-

Disease	Vector	Patients	Conditioning regime	Outcome	References
MLD	SIN LV	Italy (9)	Busulfan	On going	(Biffi, Montini et al. 2013)
WAS	γ-RV	Germany (10)	Busulfan	9/10 with benefits	(Boztug, Schmidt et al. 2010, Braun, Boztug et al. 2014)
	SIN LV	UK, USA, France, Italy (4)	Busulfan+ Fludarabine	4 with benefits	(Aiuti, Biasco et al. 2013, Castiello, Scaramuzza et al. 2015, Hacein-Bey Abina, Gaspar et al. 2015)
B-Thal	SIN LV	USA, France (4)	Busulfan	On going	(Cavazzana-Calvo, Payen et al. 2010)
		USA	Busulfan	On going	
FA	SIN LV	USA	None	On going	-
		Spain	None *	On going	-
X-ALD	SIN LV	London, France	Busulfan and Cyclophosphamide	On going	(Cartier, Hacein-Bey-Abina et al. 2009)
		USA, France (4)			

Table 2: The table summarizes the clinical trials carried in monogenic in hematological and immune diseases. ADA: Adenosine deaminase; SCID-X1: X-linked severe combined immunodeficiency; CGD: chronic granulomatous disease; SCD: sickle cell disease; MLD: metachromatic leukodystrophy; WAS: Wiskott Aldrich syndrome; B-Thal: β-thalassemia; FA: Fanconi anemia; X-ALD: X-linked adrenoleukodystrophy; SIN LV: self-inactivating lentiviral vector; RV: retroviral vector; T-ALL: T cell acute lymphoblastic leukemia; ERT: enzyme replacement therapy. Modified from Kauffman 2013 and Ghosh 2015 (Kaufmann, Buning et al. 2013, Ghosh, Thrasher et al. 2015). * The first two patients will not receive conditioning.

Gene therapy can be performed following different strategies depending on the disease to treat:

Gene substitution: It consists in the replacement of the defective gene by the corrected gene mediated or the substitution of the mutated nucleotides for the correct ones by HR.

Gene addition: A healthy copy of the defective gene, whose expression frequently is controlled by an exogenous promoter, is introduced in the target cells. This strategy can be performed either by HR or by integrative viral vectors.

Gene suppression: in order to eliminate or reduce the expression of a protein, the target gene is eliminated or reduced by knocking-out (gene targeting) or knocking-down (using mRNA interference) strategies, respectively.

• Other considerations of the Gene Therapy

- The Vector and the transgene

The vector is one of the most critical tools in GT, since it is the most frequent system employed to introduce the DNA or RNA into the cells in a safe and efficient way. In HSC diseases, the γ-retroviral (γ-RV) and lentiviral vectors (LV) are the vectors most commonly used both in preclinical and clinical studies.

The size of the therapeutic gene is another very important parameter due to not all the viral vectors are able to carry large sequences of DNA. In most cases an internal promoter is added to induce an efficient and stable expression of the transgene. Depending on the target cells, the promoter can be specifically designed. Depending on the viral vector (γ-RVs or SIN-LVs), the integration can be favored in different regions of the cell genome, and this would mediate different epigenetic modifications in the genome that could affect the expression of the transgene (Ellis 2005).

- The Target cells

The hematopoietic system is a hierarchical system in which a few number of primitive cells called hematopoietic stem cells (HSCs), are able to generate a huge number of more differentiated cells of the hematopoietic system: blood, BM, liver, spleen, thymus, lymph nodes and accessory lymphoid tissues.

These HSCs, are characterized by their capacity for self-renewing and for generating different cell types. These cells types belong to two main lineages: lymphoid, which corresponds to the adaptive and innate immune system (T, B and natural killer cells), and myeloid, responsible for the innate immune response, oxygen transport and hemostasis (granulocytes, monocytes, erythrocytes and megakaryocytes/platelets). In order to characterize the HSCs in preclinical studies, xenotransplant models consisting in the transplant of human cells into immunodeficient mice are frequently performed. These human hematopoietic precursor that are able to repopulate hematopoiesis of a SCID mouse are called SCID repopulating cells (SRCs) (Kamel-Reid and Dick 1988, McCune, Namikawa et al. 1988). The CD34 marker was defined as a reliable marker of SRCs (Civin, Trischmann et al. 1996, Larochelle, Vormoor et al. 1996). Other subpopulations have been identified with higher SRC capacity, such as the $CD34^{+}CD38^{-}$ (Bhatia, Bonnet et al. 1998, Guenechea, Gan et al. 2001). Other studies have identified additional HSC populations: cells with capacity to repopulate immunodeficient mice in the

short term (short term SRCs) or in the long term (long term SRCs) (Guenechea, Gan et al. 2001, Doulatov, Notta et al. 2012).

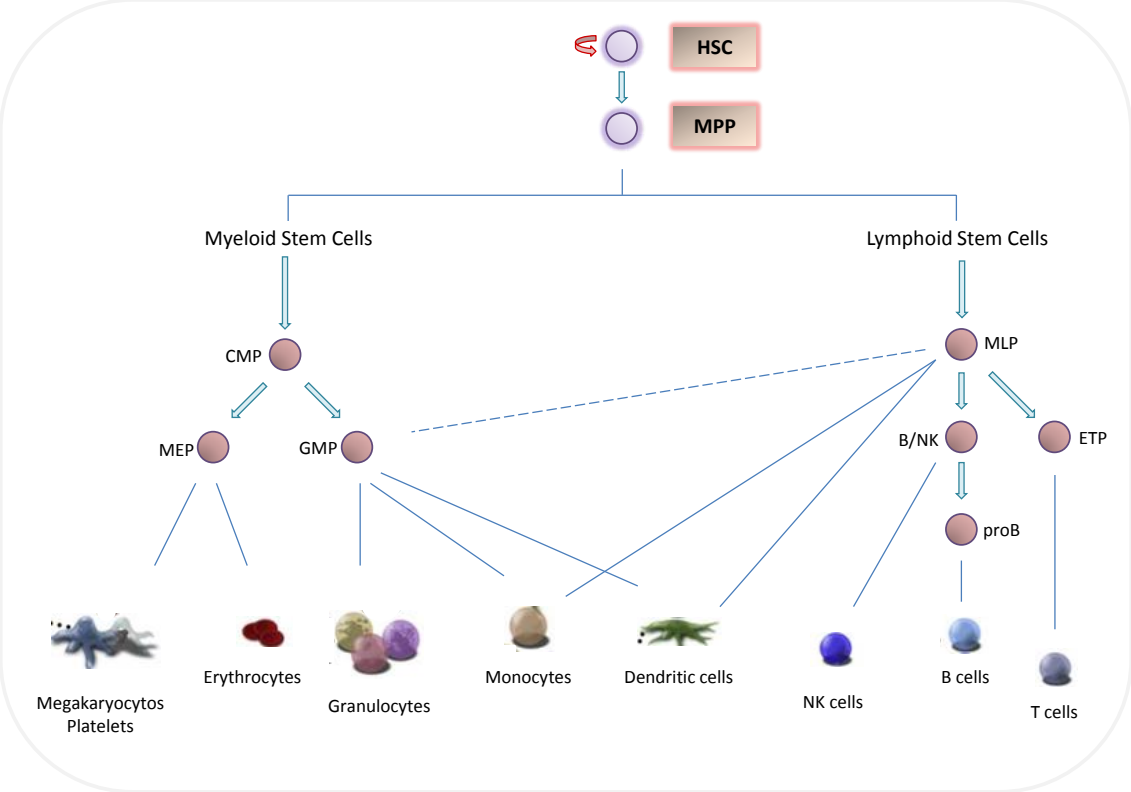

Figure 4: Current model of lineage determination in the human hematopoietic system hierarchies. Picture modified from Doulatov *et al.* (Doulatov, Notta et al. 2012).

For all the previous reasons, the ideal target cell to perform GT would be these primitive HSCs, which capacity to long term repopulate in the all lineages of the hematopoietic system.

- The efficiency of transduction and the selective advantage

The efficiency of the gene delivery is a very important parameter for the success of the GT. In this respect, the number and quantity of the HSCs to be transduced becomes critical. A low number of corrected HSCs will increase the probabilities of engrafting the patient with a recovered Hematopoietic system. For this reason, the type of vector and its transduction efficiency are critical for the success. In some cases the corrected cells acquire proliferative advantage, which means that a few precursor cells would be able to proliferate and restore the Hematopoietic system. This is for example the case of X1-SCID (Kaufmann, Buning et al. 2013) and FA (Navarro, Rio et al. 2015)

- **Immunogenicity**

This is related with the response that the immune system of the patient may exert against the transgene or certain proteins of the capsids of the viral vectors. Several studies are being performed to decrease this reaction against the proteins of the capsid, something that is particularly relevant in this GT approaches (Bessis, GarciaCozar et al. 2004, Annoni, Goudy et al. 2013, Masat, Pavani et al. 2013).

- **Insertional genotoxicity**

This is an essential factor to be considered when performing GT with integrative vectors. The effects induced by these vectors have been extensively studied in HSCs from humans and mice, due to the insertional oncogenesis effects observed in human clinical trials that can activate protoncogenes due to enhancers or aberrant splicing of the transcript of the vector (Modlich, Navarro et al. 2009, Montini, Cesana et al. 2009). For this reason, the design of the viral vector is critical, taking into account the different mechanisms of integration used by these vectors, being this pattern safer in LVs (Montini, Cesana et al. 2006, Nienhuis, Dunbar et al. 2006, Modlich, Navarro et al. 2009, Montini, Cesana et al. 2009). Both γ-RV and LV integrate in transcriptionally active areas (De Palma, Montini et al. 2005). Consequently, these vectors have been modified, with self-inactivating (SIN) Long Terminal Repeats (LTRs) (Modlich, Bohne et al. 2006, Montini, Cesana et al. 2006, Montini, Cesana et al. 2009). Furthermore, the use of weak or tissue specific promoters (Modlich, Navarro et al. 2009) and insulators (that protect surrounded genes from the vector influence) have increased the safety of these vectors.

The prolonged *ex vivo* manipulation of the HSCs and the composition of the culture media are important too, since long incubation periods limit the engraftment capacity of the HSCs (Maetzig, Brugman et al. 2011).

• Gene Therapy in Fanconi anemia

Ex vivo GT, has been proposed in FA patients based on different observations: the existence of mosaic patients, that have suffered a spontaneous mutation in the defective FA-gene in a HSC, reverts the existence of a proliferative advantage in corrected FA HSCs. So, if we manage to introduce a healthy copy of the defective gene in a limited number of HSCs, the BMF may progressively be restored. This phenomenon distinguishes FA from other monogenic diseases treated with GT, since it suggested the possibility of engraftment in the hematopoietic system of the patient even without conditioning.(Navarro, Rio et al. 2015). Additionally the good results that gene therapy has shown in the monogenic diseases strongly suggest that a similar

approach will be also efficient and safer for FA patients without an HLA-identical donor (Rio, Meza et al. 2008).

Two studies have been performed in FA patients with autologous $CD34^+$ cells transduced with γ-RVs (FA-A and FA-C patients). However, no clinical benefits were obtained from this treatments (Liu, Kim et al. 1999, Walsh, Fu et al. 2001, Kelly, Radtke et al. 2007). In one of them, three pediatric patients and one adult FA patients were treated. $CD34^+$ cells were transduced *ex vivo* and very low numbers of transduced cells (0.01-1.9 x 10^6 $CD34^+$ cells/kg body weight) were infused. Transduced cells were detected in very low numbers but no clinical benefits were observed, although neither allergic or anaphylactic reactions nor replication-competent retrovirus were observed (Walsh, Fu et al. 2001). In the other study, a first clinical trial was performed and focused in the collection of a significant number of $CD34^+$ cells, being the goal of 2 x 10^6 $CD34^+$ cells/kg future weigh in 5 years. Three of the four patients of the study achieved > 80% of cells. With respect to the GT trial, the infusion was safe and well tolerated by the patients and showed transient improvements in blood cells counts, but the BMF was not restored. The authors conclude that higher numbers of transduced $CD34^+$ cells and improved transduction rates would be necessary to achieve clinical benefit (Kelly, Radtke et al. 2007).

The critical point in FA GT that constitutes a handicap with respect to the diseases is that the number of HSCs in FA patients is significantly lower. Defects in pluripotent stem cells have been studied in FA-induced pluripotent stem cells (iPSCs) and in embryonic stem cells with downregulated FA-genes (Raya, Rodriguez-Piza et al. 2009, Tulpule, Lensch et al. 2010, Muller, Milsom et al. 2012, Yung, Tilgner et al. 2013, Suzuki, Niwa et al. 2015). A lower number of HSC have been found in BM and in PB of FA patients in comparison to healthy donors, and studies in FA mouse models have shown defects in the repopulation potential of HSCs (Rio, Navarro et al.). Due to this low number of HSCs present in the BM of FA patients, particularly in advanced stages of the disease, the collection of $CD34^+$ cells is challenging. Consequently, several studies are being performed to improve the number of HSCs, for example by collecting umbilical cord blood (UCB) samples (Navarrete and Contreras 2009) and bye mobilizing the HSCs to peripheral blood (PB) with G-CSF and perixaflor (Pulliam, Hobson et al. 2008). A clinical trial has been opened (EudraCT number 2011-006197-88) to obtain higher doses of mobilized $CD34^+$ cells in FA patients in Spain (Navarro, Rio et al. 2015).

In order to improve the transduction efficiency with the therapeutic gene, culture times have been reduced (Jacome, Navarro et al. 2006) and new vectors (SIN-LVs) have been generated (Gonzalez-Murillo, Lozano et al. 2010).

Our lab obtained the Orphan Drug designation in 2010 by the European Commission (EU 3/10/822) for a SIN LV vector developed by *Gonzalez-Murillo et al.* (Gonzalez-Murillo, Lozano et al. 2010). A Phase I/II Gene Therapy Trial for FA-A patients, the most common in Spain, (Antonio Casado, Callen et al. 2007), has been recently approved. The aim of this clinical trial is to improve the clinical efficacy of the GT in FA.

An International FA Gene Therapy Working Group was constituted in 2011 in order to set up the necessary conditions to improve the efficacy and the safety in future FA clinical trials (Tolar, Becker et al. 2012).

3. TARGETED GENE THERAPY

Although, as LV mediated GT is offering new clinical data demonstrating the safety and efficacy of this new therapy in monogenic diseases, new approaches would further improve the safety the gene therapy in the genome. Targeted GT consists in the integration, interruption or replacement of specific genes or nucleotides in the cells genome.

The main approach used for targeted gene therapy is based on the aim exploiting the natural HR DNA repair mechanism of the cell. HR is a very accurate DNA repair mechanism that uses the sister chromatid as template to repair DSBs. HR-based gene therapy strategies were started in the 1980s, using DNA donors flanked by long arms with homology to the target locus (Capecchi 1989). However, the probability of inducing the specific insertion of the exogenous sequence in the target site (genetic edition) was very low, around 10^{-6} (Cathomen and Joung 2008). For this reason several improvements have been made, the most important was the development of engineered nucleases able to recognize and to induce DSBs in a target site, with the consequent activation of the HR, and thus of gene editing (**Figure 5**).

Although the transfer of designed nucleases together with a donor template can increase the efficiency of the process, HR is not as frequent as other DSB repair mechanisms such as the non-homologous end-joining (NHEJ). This is an error prone mechanism that binds the two ends originated in the DSB, without the use of any template. If the NHEJ fixes the breakage, insertions and/or deletions can be introduced in the target site (**Figure 5**). The HR process is more frequent during the S and G_2 phases of the cell cycle, so this process takes place more frequently when cells are proliferating, and its permissiveness varies with the cell type (Ciccia and Elledge 2010).

There are several applications using both DNA repair mechanisms. If NHEJ takes place, mutations and deletions (INDELs) could be introduced in the target site. This has a particular application for the stable disruption of specific target genes (Ciccia and Elledge 2010). When homologous repair (HR) takes place, three strategies of gene editing can be approached (**Figure 5**):

1. **Gene correction**: this strategy facilitates to replace specific nucleotides mutated in a gene for the correct ones. The advantage of this method is that the expression of the wt gene is restored under its endogenous promoter, but it should be designed for each specific mutation, so it is not useful when a disease can be originated by multiple mutations.
2. **Gene targeting in a Safe Harbor (SH)**: This consist in the introduction of the cDNA of a therapeutic or reporter gene, into sequences of the genome that do not play a significant role in the cell. In many cases the expression of the transgene is controlled by an exogenous promoter. If the SH is a genomic desert the expression the transgene could be low. The main advantage is that potentially any disease originated by any mutation could be treated with the same nuclease, just modifying the therapeutic gene in the donor.

3. **Targeted Knock in insertions:** In this case the cDNA of a therapeutic gene is introduced upstream the mutated sequence of a gene. In this case the endogenous promoter will control the expression of the corrected gene. With the same nuclease all the possible mutations causing the disease in this gene, including deletions, could be potentially corrected.

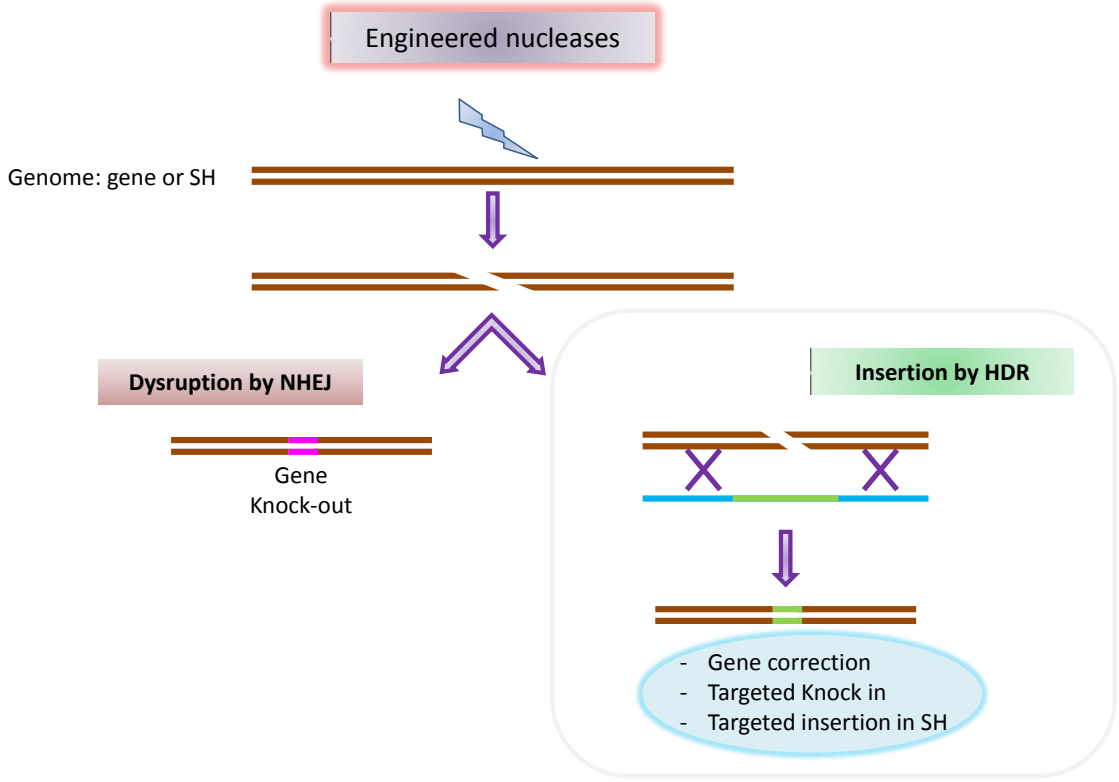

Figure 5: Scheme representing the two main mechanisms involved in the DSBs repair of engineered nucleases induce DSBs in a target site of the genome. Thereafter, if the break is repaired by NHEJ INDELs could be introduced in the genome generating gene mutations. When HR takes place with a donor template the therapeutic sequence could be integrated, inducing gene correction, or targeted insertion in specific genes (knock-in) or in Safe Harbor loci (SH).

3.1 Engineered nucleases

Although the efficiency of targeted gene editing mediated by HR is very low, around 10^{-6} events, several studies have reported that HR can be enhanced by introducing DSBs at the site of the desired recombination event (Rouet, Smih et al. 1994, Smih, Rouet et al. 1995). For these reasons important efforts have been made to set up a good platform to induce targeted DSBs, by engineered nucleases.

• Types and characteristics of the engineered nucleases

In the last years, several types of engineered nucleases have been developed due to the importance of the genetic manipulation of higher eukaryotic cells in applied biology. Engineered nucleases are artificial proteins that that contain a DNA binding domain that recognizes specific sequences in the genome, and a nuclease domain (Naldini 2011, LaFountaine, Fathe et al. 2015). Four engineered nucleases are the most frequently used in gene editing technologies:

- Zinc-Finger Nucleases (ZFN)

ZFN are artificial fusion proteins composed by a zinc finger DNA binding domain linked to a nonspecific nuclease domain (**Figure 6**). These proteins were originally named chimeric restriction enzymes and were first developed by Chandrasegaran and coworker (Kim, Cha et al. 1996, Chandrasegaran and Smith 1999).

The nonspecific nuclease domain is a type IIS nuclease domain of the restriction enzyme Fok-I (Kim, Cha et al. 1996), and its cleavage activity requires dimerization. For this reason, in order to induce the DSB in the genome, ZFNs are designed as a pair of monomers of ZFN in reverse orientation (the ZFN has DNA binding domains surrounding the target site). The DNA binding domain of each monomer is constituted by tandem repeats of C_2H_2, each of them containing a Zn^+ that interacts with a of 3 bps target in the DNA. Up to six fingers are modularly assembled, with some nucleotides separating the fingers, to generate the ZFN monomer with a specific target site (targeting 9-18 nucleotides), giving rise to a ZFN able to target 18-36 nucleotides. An amino acid linker binds the DNA binding domain with the nuclease. Both monomers, after dimerization, are separated by a "spacer" sequence (Porteus and Carroll 2005, Urnov, Rebar et al. 2010, Carroll 2011).

Figure 6: Zinc-Fingers Nucleases binding their target site in the genome and the nuclease FokI. Picture from the commercial brochure of ONA Bio.

The first experiments in which the best conformation of the ZFN was verified were conducted in the *Xenopus laevis* oocyte system (Bibikova, Carroll et al. 2001). The first genomic locus to be successfully targeted with ZFNs was the yellow gene of the fruit fly *D. melanogaster*, by

Bibikova and collaborators (Bibikova, Golic et al. 2002, Bibikova, Beumer et al. 2003). The first demonstration of gene targeting with ZFNs in human somatic cells aimed to disrupt the GFP reporter gene in HEK-293 cells (Porteus and Baltimore 2003). Since then, ZFNs have been applied in several approaches, including gene disruption, gene corrections and gene addition in mammalian somatic and embryonic cells (Urnov, Rebar et al. 2010). Genovese *et al*. were the first team performing HR experiments using ZFNs in human CD34+ cells. They manifest the difficulties of targeting these primitive cells, but finally these targeted $CD34^+$ cells from healthy donors and X-linked severe combined immunodeficiency (SCID-X1) patients were targeted by reporter and the therapeutic IL2RG gene. These targeted cells sustained normal hematopoiesis and gave rise to functional T cells with selective growth advantage compared to T cells carrying the genetic defect (Genovese, Schiroli et al. 2014). Another paper, published by Hoban et al., reported gene targeting in $CD34^+$ cells. They used ZFNs to correct the point mutation that originates the sickle-cell disease (SCD) in human CD34+ cells and HSCs. These modified cells maintained the repopulating capacity of these cells in NOD/SCID/IL2rγ^{null} (NSG) mice, originating cells of all the hematopoietic lineages in these animals (Hoban, Cost et al. 2015).

- CRISPR/Cas9 system

The CRISPR/Cas (Clustered regularly interspaced short palindromic repeats-CRISPR-associated proteins) modules are adaptive immunity systems that are present in many archaea and bacteria. They act against invading genetic elements such as viruses and plasmids (Karginov and Hannon 2010, Makarova, Haft et al. 2011). Although ten different CRISPR/Cas systems have been described so far, the type II CRISPR/Cas9 system (CRISPR/Cas9) is the most recent mechanism to induce DSB in the genome. This system is formed by the nuclease protein Cas9 which possess two catalytic domains, Ruv and HNH, that cleave the two strands in the DNA, guided by the trans-activating CRISPR RNA (tracrRNA):CRISPR RNA (crRNA) duplex (**Figure 7**). The crRNA is complementary to the target site in which the Cas9 will cleavage (Chylinski, Le Rhun et al. 2013, Gasiunas and Siksnys 2013). The dual trackRNA:crRNA was further developed as a single-guide RNA (sgRNA) for genome engineering that contains the 5´ end 20-nucleotide sequence determining the DNA target site according to Watson-Crick base pairing and 3´ end double-stranded structure binding Cas9 (Jinek, Chylinski et al. 2012). This sgRNA will guide the CRISPR-Cas9 to any DNA target sequence with an adjacent protospacer-adjacent motif (PAM) DNA sequence, composed of three bp, NGG or NAG, just changing the sgRNA sequence. The sgRNA and the Cas9 can be delivered separately, confering the CRISPR/Cas9 system a high versatility due to the possibility of targeting different loci at the same time. As happens with other designed nucleases, several studies have shown that the CRISPR/Cas9 could tolerate some mismatches between the sgRNA and the target sequence in the DNA. (Cong, Ran et al. 2013, Fu, Foden et al. 2013, Hsu, Scott et al. 2013).

Figure 7: The image obtained from http://www.addgene.org/crispr/jounglab/, shows the CRISPR7Cas9 system bound to its target DNA in the genome.

The two first works in which the CRISPR/Cas9 system was used in gene editing were published in 2013. In the first one, Mali et al. used this system to target the *AAVS1* locus in 293T, K562 cells and human iPSCs (Mali, Yang et al. 2013). Cho et al. used the CRISPR/Cas9 to target 293T cells in comparison with ZFNs and TALENs (Cho, Kim et al. 2013). Since then, more than 520 studies have been published with this system, see revision (LaFountaine, Fathe et al. 2015, Yang 2015). The CRISPR/Cas9 system has been used for the efficient genome engineering in other organisms such as *Drosophila* (Gratz, Cummings et al. 2013), *Caenorhabditis elegans* (Friedland, Tzur et al. 2013), *Axolotl* (Flowers, Timberlake et al. 2014), *Xenopus tropicalis* (Nakayama, Fish et al. 2013), rat (Hu, Chang et al. 2013), pig 26554 and cynomolgus monkey (Niu, Shen et al. 2014). The use of this system has markedly improved the generation of transgenic mice both in time and possibilities. For example the coinjection of Cas9 and sgRNAs targeting different genes into mouse zygotes have generated mutant mice with biallelic mutations (Wang, Yang et al. 2013). By the inclusion of a mutant oligo nucleotide, gene knocking mice carrying precise point mutations of two genes have been also generated (Yang, Wang et al. 2013). Additionally, by co-injecting the two-component transcriptional activator including an inactive “nuclease-dead” Cas9 protein fused to a transcriptional activation domain and sgRNAs, multiplexed activation of endogenous genes can be obtained (Cheng, Wang et al. 2013). Even a Cre-dependent Cas9 knocking mouse has been generated that could facilitate the generation of mutant mice just by injecting sgRNAs (Platt, Chen et al. 2014). Finally, CRISPR/Cas9 has been also used to the treatment of several monogenic diseases, including X1-SCD (Xie, Ye et al. 2014), Duchenne muscular dystrophy (DMD) (Li, Fujimoto et al. 2015), and coronary heart disease (Ding, Strong et al. 2014), among others.

- Meganucleases (MNs)

MNs, also called homing endonucleases (HEs), can be classified in five families based on sequence and structure motifs. The best studied family is that of the LAGLIDADG proteins, which are present in all kingdoms of life and most of them coded by introns or inteins. They are highly specific endonucleases that recognize and cleave the exon-exon junction sequence, that range from 14 to 40 bp in length, wherein they resides, a fact that gives them the name of

HE (Silva, Poirot et al. 2011). In 1997 two groups from this family highlighted, with the structure of both single-motif protein (I-CreI) (Heath, Stephens et al. 1997) and an intein-encoded double-motif protein (PI-SceI) (Duan, Gimble et al. 1997). Although several hundred of natural MNs have been identified, their targets sites do not contain all the genome sequences, making so difficult to find a MN targeting a target site of interest. For this reason, the specificity of these natural MNs has been modified by mutations in the residues of their binding domain that then are fused to create engineered MNs by a very complex process. After the engineering, MNs are studied by a functional screening. (Chen and Zhao 2005, Arnould, Chames et al. 2006, Doyon, Pattanayak et al. 2006, Silva, Belfort et al. 2006, Delacote, Perez et al. 2013). One of the positive aspects of MNs is their low toxicity in mammalian cells (Rouet, Smih et al. 1994). The extensive characterization of an engineered I-CreI derivative cleaving the human RAG1 gene demonstrated that a redesigned MN is equal to the natural I-SceI both in terms of efficacy and specificity (Grizot, Smith et al. 2009).

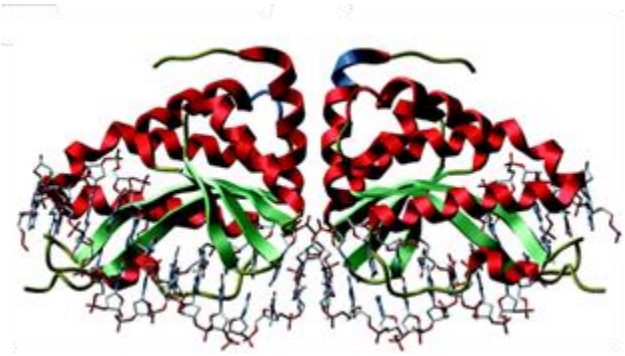

Figure 8: The picture shows the structure on an I-CreI derived Meganuclease. Picture from Epinat *et al.* (Epinat, Arnould et al. 2003).

In our work, a MN derived from the I-CreI family generated by Cellectis was designed to target a sequence inside a genomic dessert that could be considered as a SH (called SH6, which will be explained in following point) (**Figure 8**) (Eyquem, Poirot et al. 2013). In this thesis, the MNs employed are called SH6v2 and SH6v5 MNs, isoesquizomers binding the target site at the SH6 locus.

- TALE nucleases

The transcription activator-like effector nucleases (TALENs) comprise a nonspecific *FokI* nuclease domain fused to a customizable DNA-binding domain (**Figure 9**). This binding domain is composed of highly conserved repeats that derive from transcription activator-like effector (TALEs), proteins secreted by *Xanthomonas* bacteria to alter gene transcription in host plant cells (Boch and Bonas 2010). This code for TALEs DNA recognition was revealed by Adam Bogdanove (Bogdanove and Voytas 2011) and Jens Boch (Boch, Scholze et al. 2009). TALEs are modular proteins composed of an N-terminal translocation domain with central repeats that mediate sequence-specific DNA binding, and a C-terminal segment that contains nuclear localization signals and a transcriptional activator domain (Bogdanove, Schornack et al. 2010, Mussolino, Morbitzer et al. 2011). The central repeats of the TALE DNA binding domain include a variable number between 12 and 30 of conserved 33-35 residues long repeats in tandem arrays. The polymorphisms are located at positions 12 and 13 and are called repeat variable di-

residues (RVDs), which recognize one of the four bases in the target site (Boch, Scholze et al. 2009, Moscou and Bogdanove 2009, Deng, Yan et al. 2012, Mak, Bradley et al. 2012). TALENs can be easily and rapidly designed with a desired sequence specificity due to the one-to-one correspondence between the RVD modules and the bases. However, the construction of DNA segments encoding TALE arrays is difficult because the TALEN units consist of up 20 RVDs and these homologous sequences can recombine (Holkers, Maggio et al. 2013).

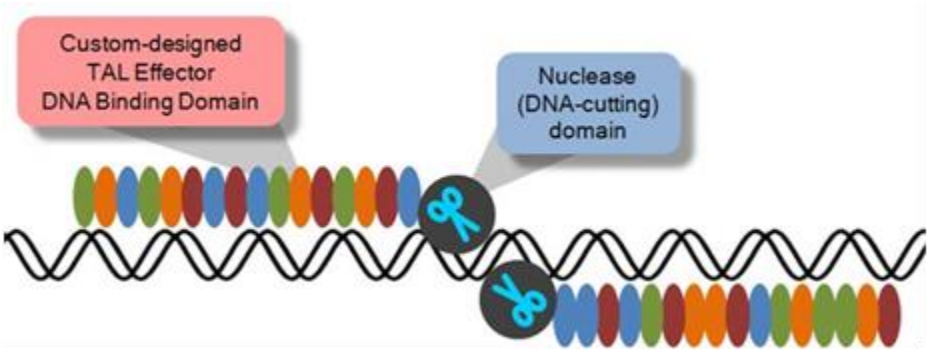

Figure 9: Each of two monomers of TALEN constitutes a heterodimer that binds to the target site in the DNA. In the picture the DNA binding domain of each monomer and the FokI nuclease are highlighted. Image from the commercial brochure of PNA Bio.

TALENs have been used for genome editing approaches in a wide range of eukaryotic systems and plants. The majority of the published works are focused on the treatment of monogenic disorders, for example: DMD. hemophilia A, (Park, Kim et al. 2014) or as β-thalassemia, (Ma, Liao et al. 2013). TALENs have also shown to specifically target mitochondrial localization sequences, in order to correct the point mutation that causes Leber´s optic neuropathy (Bacman, Williams et al. 2013). Additionally, TALENs have also been used in acquired diseases to target, for example the CCR5 gene necessary for HIV infection (Ye, Wang et al. 2014) and also the LEDGF/p75 required for HIV integration (Fadel, Morrison et al. 2014). Similarly TALENs have been used to target the viral DNA of hepatitis B virus, and the human papillomavirus, in order to limit the infection (Bloom, Ely et al. 2013, Chen, Zhang et al. 2014, Hu, Ding et al. 2015).

In this work, we have used a TALEN with a FokI nuclease domain also for the targeting of the SH6 locus. This TALEN (called SH6-TALEN) was designed by our collaborators from Cellectis.

- Next generation of engineered nucleases

In a new generation of engineered nucleases the FokI in the TALEN have been mutated, and the HNH and RuvC in the CRISPR/Cas9, catalytic domains, and have added repressors, activators or epigenetic modifiers, allowing not only the targeting of the genome, but also of the epigenome and the transcriptome (Sander and Joung 2014). TALENs with these modifications are called TALE (TAL effectors) and papers have been already published using the VP16 activation domain of herpes simplex virus, as a transactivation factor (Geibler, Scholze et

al. 2011). In the case of the CRISPR/Cas9, once the two catalytic domains of the Cas9 are inactivated, this "dead" Cas9 (dCas9) can be fused to transcriptional effector domains, such as the *trans*-repressing Kruppel-associated box (KRAB), the *trans*-activating Herpes-Simplex-Virus Protein (VP16) or the histone acetyltransferase (HAT). They have been used to down or upregulate the expression of target genes (Gilbert, Larson et al. 2013, Maeder, Linder et al. 2013, Mali, Aach et al. 2013, Perez-Pinera, Kocak et al. 2013, Hilton, D'Ippolito et al. 2015). When combined with the multiplexing gene targeting capacity of the CRISPRCas9 system, these approaches represent a valuable tool to modulate the expression of multiple genes (Konermann, Brigham et al. 2015, Mussolino, Mlambo et al. 2015).

• Considerations for the use of engineered nucleases

Before starting to work with an engineered nuclease the following respects should be carefully considered:

Success rate and activity: not all the synthesized nucleases are functional and efficient. For example, ZFNs, especially those created by modular assembly, fail to cleave the target sequences (Ramirez, Foley et al. 2008, Kim, Lee et al. 2010, Sander, Dahlborg et al. 2011). TALENs used to achieve nearly 100% of cleavage in mammalian cells (excluding hypermethylated target sites) (Kim, Kweon et al. 2013) and CRISPR/Cas9 system shows a wide range of activities in different cell types (Cong, Ran et al. 2013, Cho, Kim et al. 2013, Mali, Yang et al. 2013). Summarizing, the success rate and activity of engineered nucleases depend on the cell type and delivery method, and this can be measured example, measuring the percentage of mutagenesis (Insertions and deletions: INDELs) in the targeted locus by Surveyor assay (see Materials and Methods point 4.3).

Specificity: The specificity of ZFNs and TALENs is enhanced by the fact that these nucleases have been delivered as dimers for the cleavage of the target sequences. Generally, the size of the DNA binding domain of the nucleases enhances their specificity (Kim and Kim 2014). In theory, when a nuclease recognizes DNA sequences of at least 16 bp, the generation of off-target effects in the human genome are not common because of the complexity of a 16-bp sequences (4^{16}=4.3 x 10^9) is greater than the size of human haploid genome (3.2 x 10^9). However, all the nucleases can originate off-target mutations in mammalian genomes (Gabriel, Lombardo et al. 2011, Mussolino, Morbitzer et al. 2011, Pattanayak, Ramirez et al. 2011, Hsu, Scott et al. 2013).

Mutation signature: Each nuclease generates a different cut pattern after the cleavage, that can be recognized as a signature. Although TALENs and ZFNs, for example, share the same FokI nuclease domain, TALENs produce deletions much more frequently that insertions, while ZFNs induce insertions and deletion at similar frequencies. CRISPR/Cas9 originates frequently one- or two- nucleotide insertions (Cho, Kim et al. 2013, Kim, Kweon et al. 2013).

Delivery of designed nucleases: For the efficient genome editing, both designed nucleases and donor templates (targeting vectors, plasmids or single stranded oligo-

dendronucleotides) must to be efficiently delivered into the cells using (electroporation or liposome transfection). Generally, designed nucleases are delivered by the transfection of plasmidic DNA, or *in vitro* transcribed mRNA, or by transduction with viral vectors or purified proteins (Porteus and Baltimore 2003, Urnov, Miller et al. 2005). *In vitro* transcribed mRNA encoding the engineered nuclease and gRNAs could also be microinjected or nucleofected into cells (Hwang, Fu et al. 2013, Genovese, Schiroli et al. 2014, Poirot, Philip et al. 2015). This method is less toxic than the delivery of DNA, and leads to faster expression of the nuclease and avoids unwanted integration of plasmid DNA that encodes the nucleases. Non-integrative vectors such as integrase-deficient lentivirus vectors (IDLVs), adenovirus and adeno-associated viruses (AAVs) can also deliver the nucleases. ZFNs have been delivered in IDLVs into HSCs and embryonic stem cells (Lombardo, Genovese et al. 2007), but this kind of vectors are not compatible with TALENs because the highly homologous TALE repeats often lead to unwanted recombination events. Adenoviral vectors constitute an improved alternative for the packaging of TALENs due to their capacity to host larger DNA sequences (Holkers, Maggio et al. 2013). Proteins are rapidly degraded in the cells, limiting the activity of the nucleases to targeting and also to inducing off-target effects (Pruett-Miller, Reading et al. 2009).

Although all these considerations are required to optimize nucleases mediated gene editing, additional studies have been performed to limit the cytotoxicity of the nuclease in the target cells. This cytotoxicity can be measured by monitoring the survival of the targeted cells with increasing doses of the nuclease. Choosing a unique target site lacking highly homologous sequences in the genome is a very important first step to decrease the off-target effects by specific nuclease. Whole-exome and whole-genome sequencing analysis are a good approaches to make sure that these nucleases target unique sequences in cells (Li, Huang et al. 2011, Cho, Kim et al. 2014). One additional strategy to check if our nuclease is inducing off-target cleavage or if our donor is entering in the specific DNA sequence is to perform Southern blot assays (see Materials and Methods point 4.1), which allows to detect the presence of the expected integrations.

Despite all these precautions, most primary cells and stem cells are difficult to transfect, this implies serious difficulties for the use of engineered nucleases in these cell types. For this reason, strategies based in the use of plasmids/vectors carrying reporter genes (as fluorescent or antibiotic resistant genes) enabling the visualization and the selection of the transfected cells are frequently used (Carlson, Tan et al. 2012, Ding, Lee et al. 2013, Frank, Skryabin et al. 2013).

3.2 Gene Editing in Genomic Safe Harbors

Genomic Safe Harbors (GSH) are defined as intragenic or extragenic chromosomal locations where therapeutic transgenes can be integrated to facilitate their function in a predictable manner, without perturbing the activity of the endogenous or surrounding genes (Sadelain, Papapetrou et al. 2012). SH should obey three main criteria (Sadelain, Papapetrou et al. 2012):

- SHs should facilitate sufficient transgene expression, to yield desired levels of the foreign protein or non-coding RNA.

- SH should not to predispose targeted cells to a malignant transformation, or to alterations of their cellular functions.

- The outcome of cells targeted in SHs must to be predictable, an observation that is based on the prior knowledge.

When SHs are located in selected genes, it is assumed that their disruption of these genes is not pathological. Some studies have reported that intragenic regions could be a good choice to host expression cassettes without detectable consequences for the cell (Lombardo, Cesana et al. 2011). Three sites in the in the human genome have been stablished as SHs:

- *Adeno-associated virus site 1 (AAVS1)*: It is located in chromosome 19 (position 19q13.42) and was identified as a repeatedly recovered site of integration of wt AAV in the genome of cultured human cells that have been infected with AAV *in vitro* (Kotin, Linden et al. 1992). Integrations in this locus disrupt the gene phosphatase 1 regulatory subunit 12C (*PPP1R12C*, also known as *MBS85*) that encodes a protein whose function has not been yet elucidated. A robust expression of reporter transgenes integrated in this locus has been reported in embryonic stem cells (ESCs) and iPS cells (Smith, Maguire et al. 2008, Zou, Sweeney et al. 2011). The expression obtained across several different cell types relates to a DNAse I hypersensitive site and an insulator element contained in this site that maintains the chromatin open. The *AAVS1* locus is surrounded by several genes whose activity is not significantly modified by regulatory sequences inserted into this locus (Lombardo, Cesana et al. 2011).

- *CCR5*: Is located on chromosome 3 (position 3p21.31) and encodes the major co-receptor for HIV-1. It was defined as a GSH since the discovery that people harboring a null mutation (CCR5Δ32) in this gene were HIV resistant and did not suffer any major pathology, (Liu, Paxton et al. 1996, Perez, Wang et al. 2008). The expression of integrated transgenes is high in T-lymphocytes, monocytes and macrophages, but not in B lymphocytes or dendritic cells, and is also expressed in neurons, microglia, endothelium and smooth muscle (Rottman, Ganley et al. 1997). Although no marked abnormalities have been detected, studies in humans have demonstrated, by the KO of *CCR5,* an association between the knock-out of *CCR5* and the susceptibility of infection by the West Nile virus (Lim, Glass et al. 2006). GFP reporter genes have been integrated in this locus in cord blood CD34+ hematopoietic progenitors, T cells and human ESC lines (Lombardo, Genovese et al. 2007, Lombardo, Cesana et al. 2011) and its expression was lower than that observed in the AAVS1 in T cells (Lombardo, Cesana et al. 2011). The *CCR5* locus is located in a gene-rich genomic site, as the *AAVS1* SH, including cancer related genes that could be dysregulated by integrated transgenes (Lombardo, Cesana et al. 2011)

- *Human ROSA26*: This gene is located on the chromosome 3 (position 3p25.3) and was discovered by Irion *et al.* These authors integrated a red fluorescent protein in order to elucidate if this gene was a SH. These authors discovered that the expression of the transgene was maintained in multiple adult human tissues at variable levels, and showed no alterations, in targeted cells. However, ROSA26 is also surrounded by several genes that could be dysregulated by transgene targeting (Irion, Luche et al. 2007).

Another SH alternatives are extragenic locations of the cell genome. In this case the expression of the inserted gene is more problematic, although the integration is safer due to evolutionary considerations.

The genomic locations of endogenous retroviruses give us some clues to identify SHs. Around 8% of the human genome comprises fragments of retroviruses integrated in the germ line of mammalians during evolution (International Human Genome Sequencing, Adekoya et al. 2001). As most of them are fixed in the human germ line, we can deduce that they are not harmful, and because they are not randomly distributed but rather enriched outside of transcription units, the evolutionary selection would have eliminated their integration in transcription sites (Smit 1999).

There are five criteria to facilitate the recognition of an extragenic SH to limit risks of disruption of endogenous coding genes and ultra-conserved regions, and thus to reduce the possibility of long-range interactions between foreign transcriptional activators and adjacent genes. These criteria are the following (Bejerano, Pheasant et al. 2004, Papapetrou, Lee et al. 2010):

1. Distance of more than 50 kb from 5´ end of any gene
2. Distance of more than 300 kb from cancer-related genes
3. Distance of more than 300 kb from any microRNA
4. Outside a gene transcript unit
5. Outside of ultra-conserved regions

In this study we have worked with a novel SH, called **Safe Harbor 6 (SH6)**, characterized by fitting the five criteria exposed above, and to be a genomic desert. SH6 is located in an intergenic region in chromosome 21 and the closest ORF (C21orf37) is located more than 350 kb away (**Figure 10**). SH6 was first studied by Eyquem *et al.*, whose results suggested that it could be an useful loci for transgene expression, although the transgene expression is lower than in other intergenic SHs due to the less permissive epigenetic status of the chromatin (Eyquem, Poirot et al. 2013).

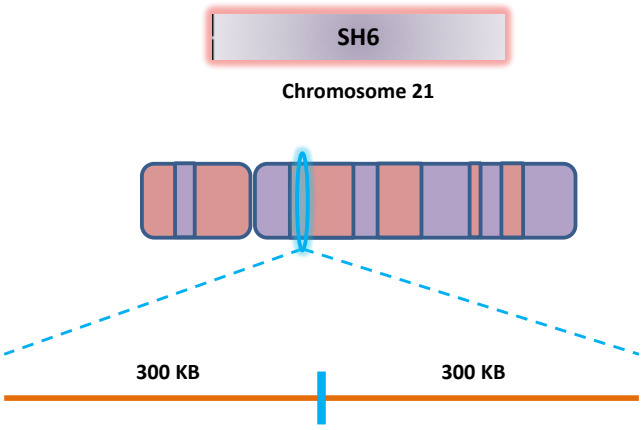

Figure 10: The picture shows the location of SH6 in chromosome 21, with 300 kb of human genome at both sides of the target site. Image modified from Eyquem et al. (Eyquem, Poirot et al. 2013).

3.3 Gene Editing and Fanconi anemia

Although the use of SIN-LVs has improved the hematopoietic GT in clinics, risks of insertional oncogenesis cannot be completely avoided due to their preferential integration into genes (Kaufmann, Buning et al. 2013). For this reason, gene editing becomes an important alternative due to their capacity to target specific sites in the genome, limiting insertional oncogenesis risks. Furthermore, gene editing makes easier the integration of the therapeutic gene into the homologous gene (knock-in or gene correction) or into heterologous genes (SH strategy) (**Figure 11**). This last possibility, is of particular interest in the case of FA taking into account that FA cells could present multiple different mutations in any of 19 FA genes (Castella, Pujol et al. 2011).

A few studies have already demonstrated that gene editing is possible in FA-A cells, even though these cells have limiting defects in DNA repair.

ZFNs have been used to target the *AAVS1* locus in combination with an integration-defective LV that carried *FANCA* gene flanked by arms with homology to the *AAVS1* locus. Targeted FA-A fibroblasts showed proliferative advantage after a few weeks of *in vitro* culture (Rio, Banos et al. 2014). Additionally, P53-downregulated iPSCs from FA-A patients with a point mutation in the fourth exon of *FANCA* were targeted by the replacement of this exon with the respective wt version, using helper-dependent adenoviruses (Liu, Suzuki et al. 2014). Finally, conventional CRISPR/Cas9 nucleases and a nickase version of this system (that only generates single-strand breaks) have shown that the nickase presents a higher efficiency to mediate DNA repair by HDR in the *FANCC* gene, as compared to NHEJ events (Osborn, Gabriel et al. 2015).

A careful design and deep analyses are necessary in the design of these engineered nucleases particularly for the targeted GT of FA cells, because of the off-target consequences of the generation of spontaneous DSBs in this DNA-repair deficient cells (Navarro, Rio et al. 2015, Osborn, Gabriel et al. 2015).

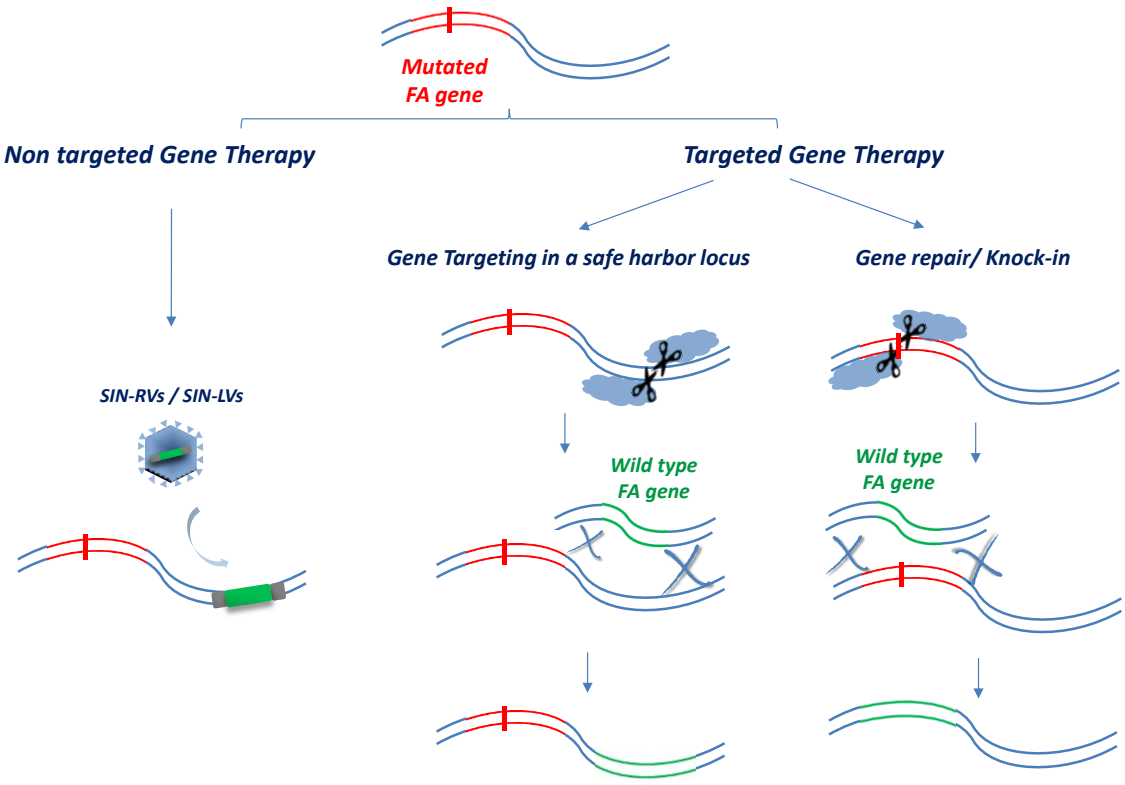

Figure 11: Different approaches followed to perform gene therapy in FA. The left panel shows the traditional gene therapy approach in which SIN-RVs or SIN-LVs carrying the FA genes are employed. In the right panel, are represented the two strategies performed in gene editing: in heterologous genomic regions, using the SHs, or in homologous regions performing Knock-in. The scheme was obtained from (Navarro, Rio et al. 2015).

OBJECTIVES

In this Thesis we aimed to develop a gene therapy strategy based on the specific insertion of foreign genes into a new "safe harbor" site (SH6), with the final goal of correcting the hematopoietic phenotype of Fanconi anemia patients.

To achieve this main objective, we focused our studies in two principal goals:

1. To investigate the efficacy and specificity of gene targeting in the SH6 "safe harbor" locus of HEK-293H cells using different meganucleases and TALE nucleases, together with a donor construct harboring *EGFP* or *FANCA* genes.

2. To develop an experimental protocol to perform gene targeting into the SH6 site of human hematopoietic progenitor stem cells using specific meganucleases and TALE nucleases, together with a donor construct harboring the *EGFP* marker gene.

MATERIALS AND METHODS

1. CELLULAR CULTURES

1.1 Adherent cells culture: HEK-293H

HEK-293-H cells (Gibco 11631-017) were cloned from the original 293 cell line and adapted to Gibco® CD 293 Medium to grow in serum-free medium. The HEK-293H cell line is a permanent line established from primary embryonic human kidney and transformed with sheared human adenovirus type 5 DNA and presents good adherence during plaque assays, superior transfection efficiencies and a high level of protein expression. The E1A adenovirus gene is expressed in these cells and participates in transactivation of some viral promoters, allowing these cells to produce very high levels of protein.

Optimal transfection efficiencies are obtained with Lipofectamine™ 2000 or 293fectin™ transfection reagents.

https://www.thermofisher.com/order/catalog/product/11631017?ICID=search-product

• Culture conditions

DMEM (Dulbecco's Modified Eagle Medium) with GlutaMAX (Gibco/Life Technologies/Thermo Fisher Scientific, Waltham, USA)) + 10% Hyclone (GE Healthcare) + 1% Penicillin/Streptomycin (P/S, Gibco). Cells were growth in adherence in normoxia: 37°C, 21% O_2, 5% CO_2 and 95% relative humidity (RH) in 75Flask.

1.2 Primary cells culture: $CD34^+$ cells derived from umbilical cord blood

• Purification of $CD34^+$ cells

Umbilical cord blood (UCB) samples from healthy donors were obtained from the Centro de Transfusiones de la Comunidad de Madrid, according to the protocol approved by the Ethical Committee. Mononuclear cells were purified by Ficoll-Paque PLUS (GE Healthcare, Fairfield, USA) density by gradient centrifugation. $CD34^+$ cells were then selected using CD34 MicroBead Kit. Magnetic-labeled cells were isolated with a LS column in QuadroMACS™ separator (MACS, Miltenyi Biotec, Bergisch Gladbach, Germany) following manufacturer's instructions. Purified $CD34^+$ were then evaluated for their purity by flow cytometry. Purities from 95-99% were routinely obtained.

• Culture conditions

- **Pre-stimulaion media:** Stem Spam (StemCell Technologies, Vancouver (Canada) + 1% P/S + 100 ng/mL stem cell factor (SCF, EuroBiosciences, Friesoythe, Germany) + 100 ng/mL FMS-like tyrosine kinase 3 ligand (Flt3, EuroBiosciences) + 100 ng/mL thrombopoietin (TPO, EuroBiosciences).

- **Culture media:** Stem Spam + 1% P/S + 300 ng/mL SCF + 300 ng/mL Flt3 + 300 ng/mL TPO.

Cells were grown in suspension in tissue non-treated 96 wells and 24 wells plates (Gibco), at 10^6 cells/mL in normoxia: 37°C, 21% O_2, 5% CO_2 and 95% RH.

2. PLASMIDS

The engineered nucleases employed in this work were two Meganucleases (MNs) and one TALEN called, synthesized by our collaborators from Cellectis.

All enzymes used during the cloning such as restriction enzymes, ligases, kinases and phosphatases were obtained from New England Biolabs (Ipswich, USA). After each ligation, the resulting products were transformed in TOP10 bacteria (Invitrogen). Plasmid DNA was purified from selected colonies or from the restriction mixes using NucleoBond® Xtra Midi EF kit (Macherey-Nagel, Duren, Germany) and analyzed by restriction enzymes, and the PCRs were performed using the Herculase II Fusion Enzyme (Agilent Technologies).

2.1 Engineered nucleases plasmids

The three nucleases were delivered into cells in plasmids carrying the sequences codifying for the nucleases, which expression was driven by the cytomegalovirus promoter (CMV). Bacterial plasmids codifying for the engineered nucleases were generated using typical procedures of molecular biology.

• MNs plasmids

Two plasmids encoding for the SH6v2 or the SH6v5 I-CreI derived MNs isoesquizomers, targeting the TTAATACCCCGTACCTAATATTGC sequence in the *SH6* locus, were used in this work. MNs are constituted by two focused monomer that are fused and transcribed by a unique plasmid. These plasmids were used in the entire gene editing experiments in which the MNs are delivered as DNA.

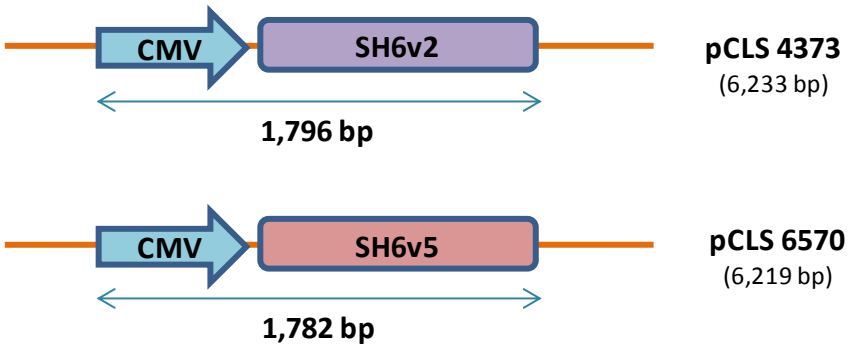

Figure 12: The picture shows the two plasmids codifying for the MNs under the CMV promoter. CMV: cytomegalovirus promoter; SH6: Safe Harbor 6; v2: version 2; v5: version 5. The blue arrow under each scheme represents the size of both the MN and the promoter. In the right side, the pCLS number represents the name of the plasmid with the size of the whole plasmid carrying the MN.

The SH6v5 MN was also delivered as mRNA, synthesized *in vitro* (see Materials and Methods section 3), in some gene editing experiments. For that reason, the sequence coding for the MN was amplified by PCR and cloned into the plasmid of the Zero Blunt PCR Cloning Kit, (Invitrogen, life Technologies), and checked by restriction enzymes for the correct insertion orientation. The forward primer contained a T7 promoter, necessary for the mRNA synthesis.

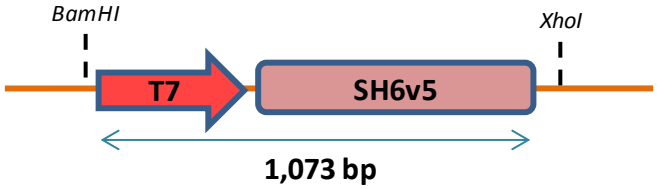

Figure 13: The picture shows the amplified SH6v5 MN amplification with the primers containing the T7 promoter and cloned into the cloning plasmid of the Zero Blunt, preceded by the T7 promoter necessary for the mRNA synthesis and surrounded by the BamHI and XhoI restriction sites used for the mRNA synthesis. The blue arrow under the scheme represents the size of both the MN and the promoter

Primer	Sequence
mRNA MN 1F	**TAATACGACTCACTATAGGGAGA**GCCACCATGGCCAATACCAAATATAACG
mRNA MN 1R	CTAAGGAGAGGACTTTTTCT

Table 3: Name and sequence of the primers used in to clone the sequence of the SH6v5 MN into the Zero Blunt PCR cloning plasmid. In bold the sequence of the T7 promoter included in the forward primer.

• TALEN plasmids

TALEN subunits, the left or the right, targeting *SH6* locus are contained in the following two plasmids provided by Cellectis. Each of these plasmids contains the sequence codifying for one of the subunits, the left or the right, of the DNA binding domain sequence of the SH6-TALEN to the SH6 locus TCTAAAGATTAATACCCCGTACCTAATATTGCATTTCCTTCTACCAGTA sequence and a FokI nuclease. The TALEN will only perform the cleavage after dimerizing with the other FokI subunit of the other plasmid.

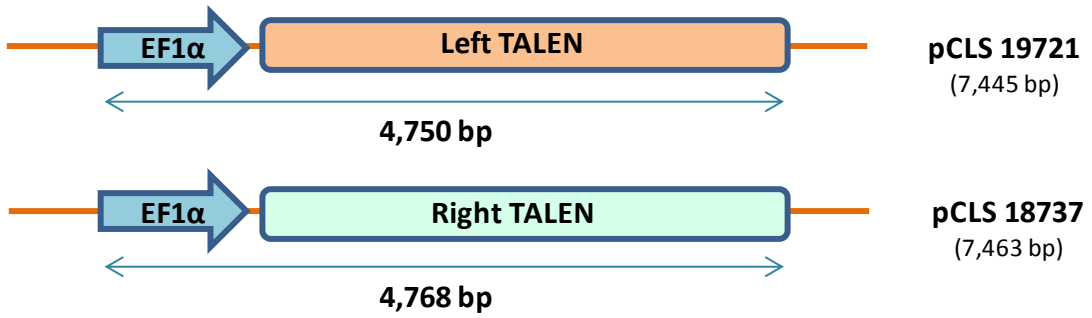

Figure 14: The picture shows the plasmids carrying the sequences codifying for the two subunits of the TALEN targeting the SH6 locus, under the control of the EF1α promoter. The blue arrow under each scheme represents the size of both the TALEN and the promoter. In the right side, the pCLS number represents the name of the plasmid with the size of the whole plasmid carrying the TALEN.

This TALEN was also delivered as mRNA, with or without the 3` UTR sequence from the β-globin, synthesized *in vitro* (see Materials and Methods section 3), in some of the gene targeting experiments. To achieve this aim, our collaborators from Cellectis gave us another pair of plasmids carrying the sequences codifying the left or the right subunits of the DNA binding domain sequence of the SH6-TALEN to the *SH6* locus TCTAAAGATTAATACCCCGTACCTAATATTGCATTTCCTTCTACCAGTA sequence and a FokI nuclease. The expression of this sequence was controlled by the T7 promoter, as with the SH6v5 MN previously commented, necessary for the mRNA synthesis.

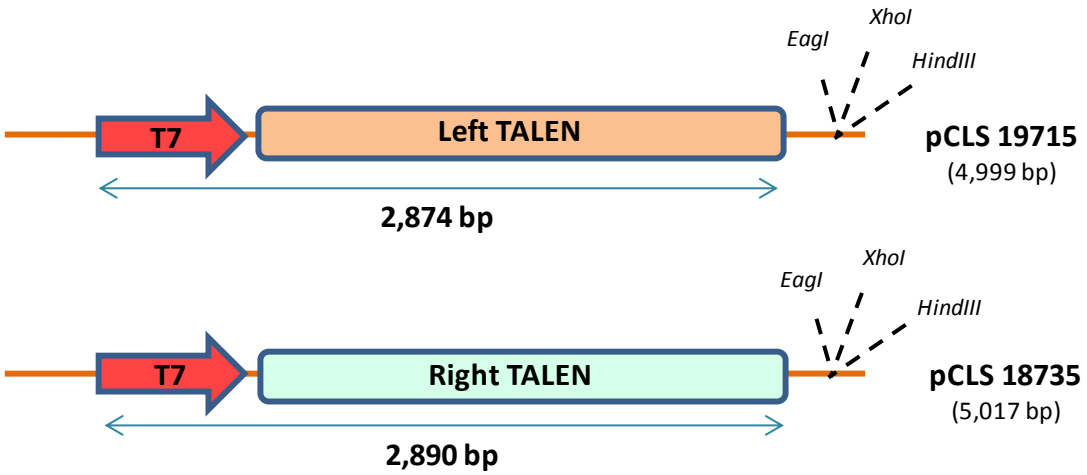

Figure 15: The diagram represents the two plasmids carrying the sequences of the right or left subunit of the SH6-TALEN targeting the SH6 locus, both of them under the control of the T7 promoter. After the TALEN sequence, the *HindIII* and *EagI*/*XhoI* restriction sites used for the mRNA synthesis and the cloning of the 3´UTR, respectively, are represented in italic. The blue arrow under each scheme represents the size of both the TALEN and the promoter. In the right side, the pCLS number represents the name of the plasmid with the size of the whole plasmid carrying the TALEN.

In order to clone the 3` UTR structure from the β-globin into the plasmids shown in **Figure 15**, two different primers (3´UTR F and 3´UTR R) shown in **Table 4** were used. Equal quantities of these primers at 10 µM were combined to be annealed by PCR with Herculase in 50µl reaction, 30 cycles (95° 15 s; 60° 20 s; 72° 25 s). After the PCR, the insert was purified by column kit and digested with *EagI*/*SalI* enzymes. Plasmids shown in **Figure 15** were digested with *EagI*/*XhoI* enzymes and dephosphorylated. Both digestions were ligated with the T4 ligase and Top10 bacteria were transformed. The colonies were checked by PCR, and the selected ones checked by restriction enzymes and sequenced with the primers shown in **Table 4** (TALEN SQ 3´UTR F, TALEN SQ 3´UTR 2 and TALEN SQ 3´UTR F 2R).

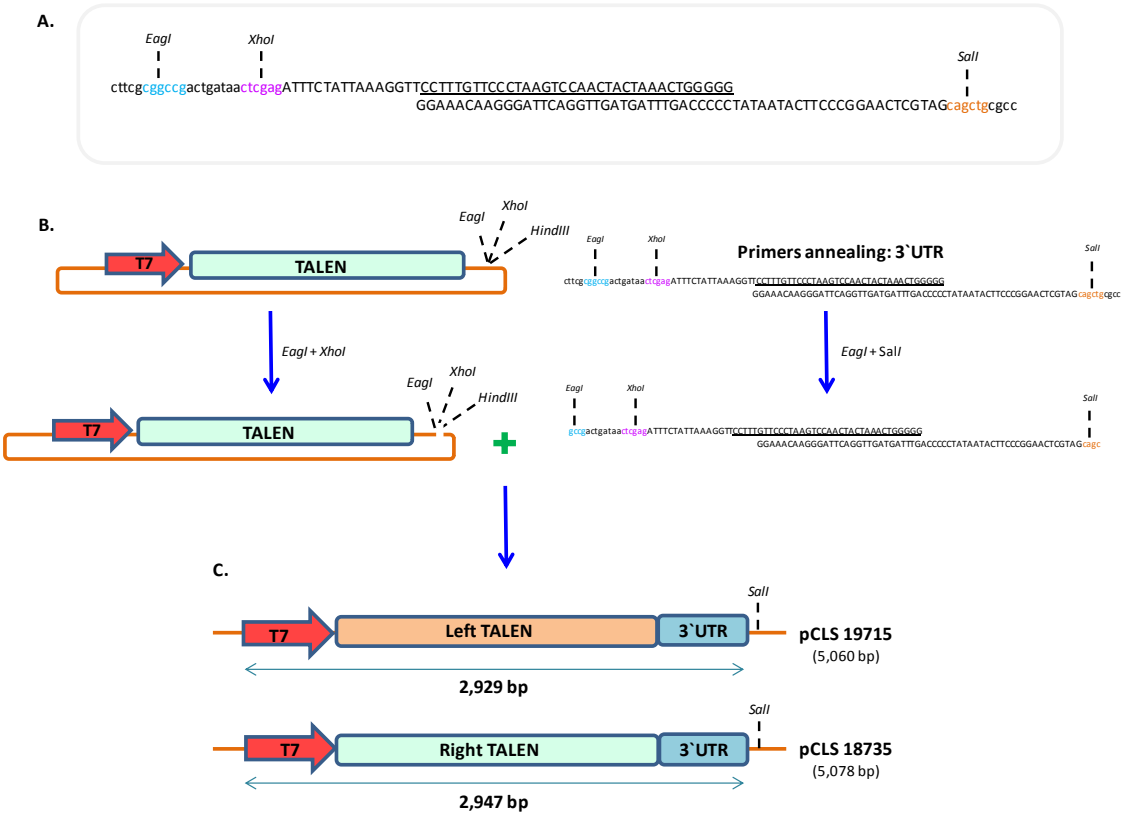

Figure 16: **A.** The pair of primers used to clone the 3`UTR is represented in the top of the figure. In upper case is the 3´UTR, in lower case are the restriction enzymes sequences and extra bases. The underlined sequences are complementary. **B**. The annealing of the 3´UTR plasmids and the plasmids containing the TALEN are digested with the corresponding enzymes and ligated. **C.** The final plasmids carrying the 3´UTR are represented in the bottom. Note that the *XhoI* restriction site is transformed into a *SalI* site after the cloning. The blue arrow under each scheme represents the size of both the TALEN and the promoter. In the right side, the pCLS number represents the name of the plasmid before the cloning process with the size of the whole plasmid carrying the TALEN.

2.2 Donor plasmids

The donor templates employed in this work were also plasmids that were co-delivered with the previous ones plasmids carrying the nucleases in the gene editing experiments in order to be used as template by the DNA repair mechanisms of the cells to fix the DSB. These donor plasmids must contain two sequences, called HAs, which are homologous to the *SH6* locus which is targeted by the nucleases. The sequence of these HAs was designed by our collaborators from Cellectis.

- ## Green Matrix (GM) or Green donor

The GM carries an EGFP reporter gene which expression is controlled by the long EF1α promoter, both of them flanked by 1.5 kb homology arms to the *SH6* locus. This donor has been used in gene editing experiments in both cell types, HEK-293H and $CD34^{+}$ cells.

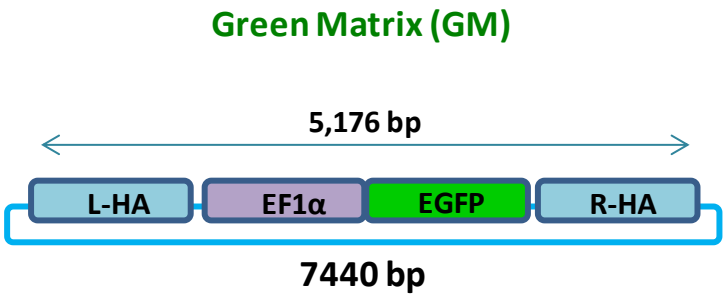

Figure 17: Picture represents the diagram of the GM donor of 7,440 bp. L-HA: Left homology arm; R-HA: Right homology arm; EF1α: Human elongation factor-1 alpha. The blue arrow under each scheme represents the size of both the TALEN and the promoter.

- ## Therapeutic Matrix (TM) or Therapeutic donor

The TM carries the WT cDNA sequence of the *FANCA* gene with a 3xFlag signaling peptide in the 5` N-termini of the protein, a self-cleaving 2A peptide (E2A) sequence (Szymczak, Workman et al. 2004) and a puromycin resistant gene as selectable marker. The long EF1α promoter controls the expression of both previous genes. The four elements are flanked by 1.5 kb homology arms to the *SH6* locus. This donor has been used in gene editing experiments in HEK-293H cells.

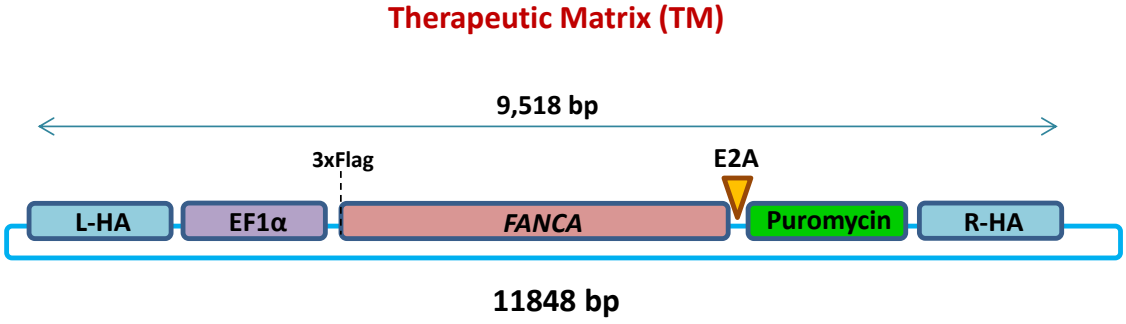

Figure 18: Diagram of the TM of 11848 bp. L-HA: Left homology arm; R-HA: Right homology arm; EF1α: Human elongation factor-1 alpha; E2A: self-cleaving peptide E sequence.

The 3xFlag-FANCA-2A-Puro construct was chemically synthesized by GenScript (Piscataway Township, USA) in a pUC57 backbone. The 3xFlag signaling peptide added to the N-termini of the protein is reported not to produce any alteration in the protein function or localization in this position (Yagasaki, Adachi et al. 2001).

The cloning strategy is represented in the following **Figure 19**.

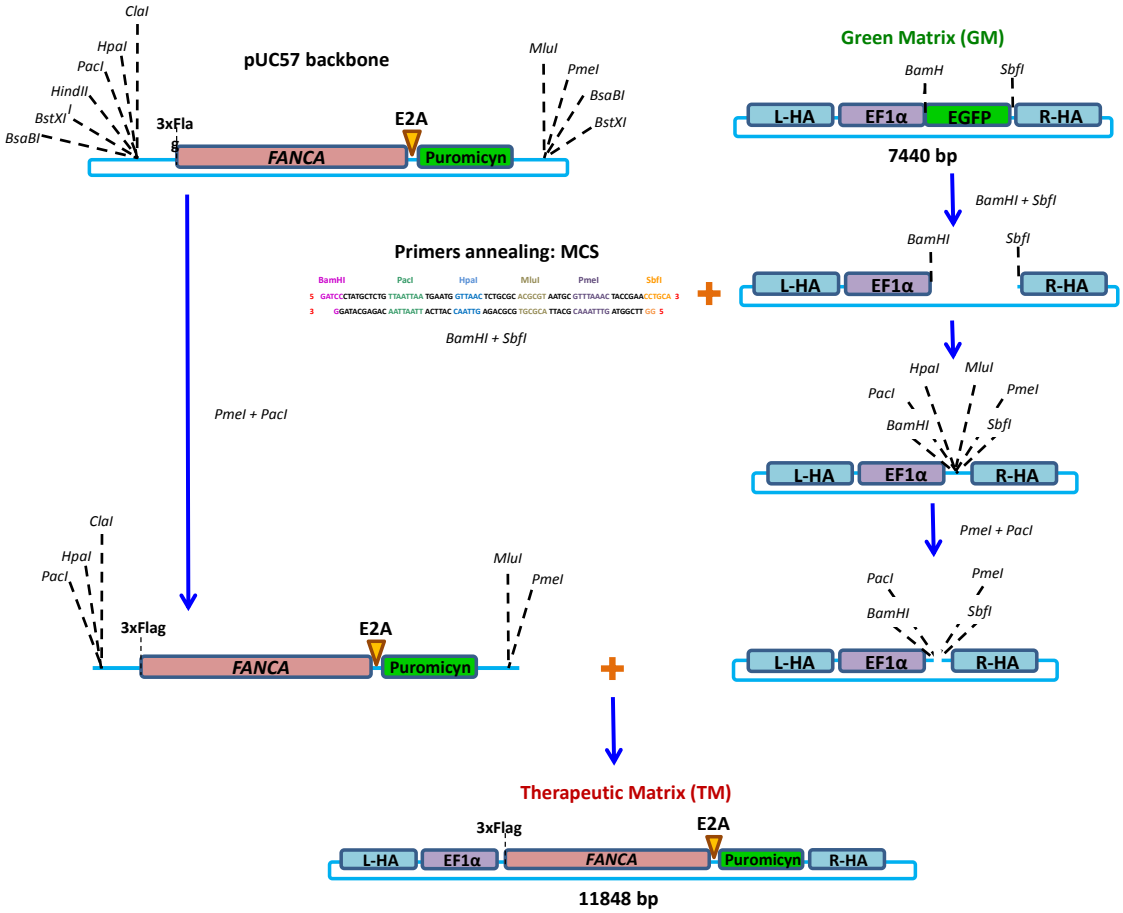

Figure 19: Diagram of the cloning strategy to build the Therapeutic Matrix introducing the construct synthesized by GenScript between the two HA to the SH6 site.

At first, a multicloning site was introduced in the plasmid: *BamHI-PacI-HpaI-MluI-PmeI-SbfI*.

BamHI	PacI	HpaI	MluI	PmeI	SbfI
5 GATCCCTATGCTCTG	TTAATTAA	TGAATG	GTTAAC	TCTGCGC	ACGCGT AATGC GTTTAAAC TACCGAACCTGCA 3
3 GGATACGAGAC	AATTAATT	ACTTAC	CAATTG	AGACGCG	TGCGCA TTACG CAAATTTG ATGGCTT GG 5

Figure 20: Primer annealing performed in our cloning strategy to introduce a multicloning site containing the *BamHI* and *SbfI* enzymes restriction sites. In different colors are highlighted the different restriction sites for the different enzymes.

Two different primers shown in **Table 4** (MCS FANCA sint 1F and MCS FANCA Sint 1R) were designed as shown in **Figure 20**. Equal quantities of these primers at 100 µM were combined to be annealed by incubation during 4 minutes at 95°C followed by 10 minutes at 70°C in annealing buffer (100 mM potassium acetate, 30 mM HEPES pH 7.4 and 2 mM magnesium acetate). The mixture was cooled down at room temperature (RT) and phosphorylated using a PNK kinase in a T4 DNA ligase buffer during 30 minutes at 37°C, followed by 10 minutes at 70°C.

At the same time the GM plasmid was digested with *BamHI* and *Sbf*I restriction enzymes in order to eliminate the EGFP reporter gene. The annealed primers and the resulting purified digestion from the GM were ligated with a T4 DNA ligase, transformed in TOP10 bacteria. Colonies were checked by PCR, as shown in **Figure 21**, using the primers FANCA Sint 2F and FANCA Sint R, to isolate the ones carrying the first intermediate of the cloning (size 6,622 bp, shown in **Figure 19**) and sequenced, with the same primers used in the PCR to choose the best one (**Table 4**).

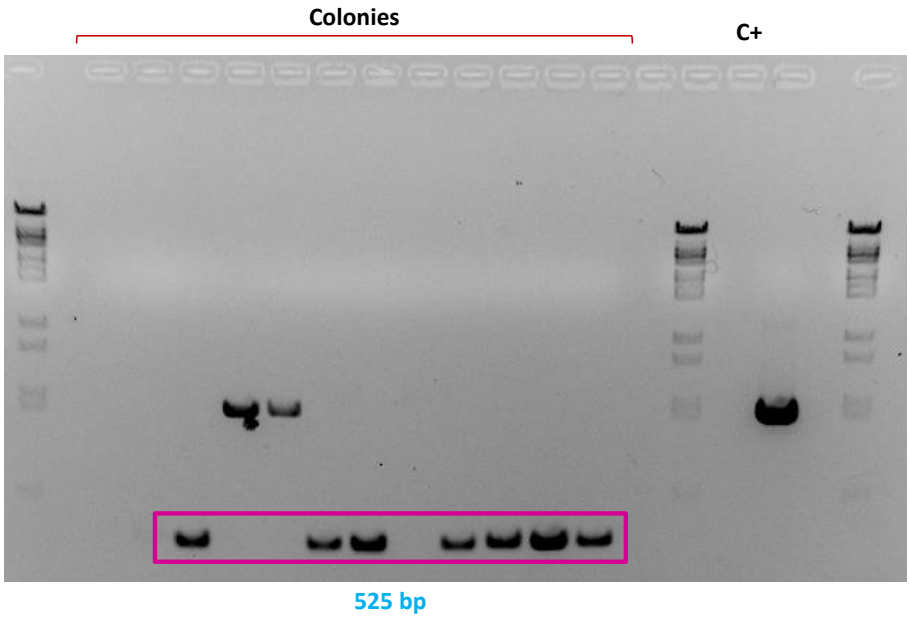

Figure 21: Electrophoresis in agarose gel showing the result test performed by PCR of the colonies after the ligation of the annealed primers and the plasmid. The bands of 525 bp inside the purple square represent colonies carrying the first intermediate of the cloning.

Then the first intermediate and the pUC57 backbone carrying the 3xFlag-FANCA-2A-Puro construct were digested by *PmeI* and *PacI* restriction enzymes. The first intermediate and the dephosphorylated 3xFlag-FANCA-2A-Puro construct (using the Antartica dephosphatase) were ligated with a T4 DNA ligase, transformed in TOP10 bacteria and colonies were checked by PCR to isolate the ones carrying the final construct of the cloning (11,848 bp) (**Figure 18**). The positive colonies analyzed using the primers FANCA Sint 2F and C+ FANCA Sint 2R (**Table 4**) that showed a band of 2293 bp where checked by restriction enzymes and sequenced with the same primers.

Primer	Sequence
3´UTR F	cttcgcggccgactgataactcgagATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGG
3´UTR R	ccgcgtcgacGATGCTCAAGGCCCTTCATAATATCCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGG
MCS FANCA Sint 1F	GATCCCTATGCTCTGTTAATTAATGAATGGTTAACTCTGCGCACGCGTAATGCGTTTAAACTACCGAACCTGCA
MCS FANCA Sint 1R	GGTTCGGTAGTTTAAACGCATTACGCGTGCGCAGAGTTAACCATTCATTAATTAACAGAGCATAGG
FANCA Sint 2F	ACCCACACAAAGGAAAAGGG
FANCA Sint R	GAATGGTCTGAAAATGGAGAGGTTAAATGAGATTT
FANCA Sint 2R	CCACTCTCTCTGCATCTGAA
TALEN SQ 3´UTR F	GGAGGAAGTTCAACAACGGCGAGATCAA
TALEN SQ 3´UTR 2F	CATCAACCCCAACGAGTGGTGGAA
TALEN SQ 3´UTR R2	TGGCCTTTTGCTGGCCTTTTGC

Table 4: Name and sequence of the primers used to clone the 3`UTR TALEN and the Therapeutic Matrix.

3. mRNA SYNTHESIS

In order to be linearized, T7-MN and T7-TALEN plasmids were digested with *BamHI* and *XhoI* or with *HindIII* restriction enzymes, respectively. This digestion was purified using sodium acetate buffer solution, pH5.2 (Sigma-Aldrich) and ethanol. The amount of DNA was measured using a Nanodrop®. The *in vitro* synthesis of the mRNA was performed with the mMESSAGE mMACHINE® T7 Ultra Kit, Synthesis of Translation Enhanced Capped Transcripts (Ambion/Life Technologies/Thermo Fisher Scientific, Waltham, USA). Firstly, 1 µg of the digested and purified product was used to synthesize mRNA following the vendors' guidelines and was purified by LiCl precipitation. To obtain a proper quantity of mRNA, 3 reactions were performed for each plasmid. For this reason the quantities of each reactive during the protocol is 3 times higher. Finally, the amount of mRNA was measured with Nanodrop®, aliquoted and stored at -80°C.

In order to test the quality of the synthesized mRNA before and after the polyadenylation reaction, an aliquot was run in an agarose gel in denaturing conditions (with formaldehyde).

Each *in vitro* nuclease mRNA production, either the mRNA MN or the TALEN pair, was tested in $CD34^+$ cells using the AMAXA Nucleofector kit for human $CD34^+$ cells (Lonza). Forty eight hours after the nucleofection, cell survival was tested and the activity of the mRNA for each nuclease production was evaluated by Surveyor assay (see Materials and methods point 4.3), in order to determine the optimal concentration of nuclease to be used in the different productions.

4. HOMOLOGOUS RECOMBINATION EXPERIMENTS

Gene targeting experiments were performed in two different types of cells described in the previous point (1. CELLULAR CULTURES of Materials and Methods). The protocol followed for these experiments are described in the following sections.

4.1 Gene targeting in HEK-293H cells

HEK-293H cell line (Invitrogen) was cultured as described in Materials and Methods point 1. One million and a half of these cells were seeded in p100 plates and the next day, plates carrying cells with a confluence around 80% were transfected with Lipofectamine 2000 (Invitrogen) following vendor guidelines. Different amounts of the MNs or the TALEN (see RESULTS points 1.2, 2.1 and 3.2) were co-transfected with 4 µg of the Green Matrix (GM) or the Therapeutic Matrix (TM) donor plasmids. Three days post-transfection, cells of each condition were harvested, counted and plated in 2 p96 plates (10 cells per well). The rest of the cells were maintained in culture in the same conditions and their EGFP expression was measured by Flow Cytometry at different time points up to 30 days (see Materials and Methods following point of Flow cytometry analysis in HEK-293H cells).

P96 plates were duplicated after 7 days in culture and cultured until the wells were confluent. At this point, the duplicated plates were frozen in viability. This allowed us to establish and expand pseudoclones in the future (see Freezing/Thawing of HEK-293H cells and pseudoclones establishments point in Materials and Methods). The other plates were maintained to extract the gDNA of the cells (see point Materials and Methods, gDNA extraction of the HEK-293H cells in p96 plates) in order to perform specific targeted analysis (see point in Materials and Methods of Specific gene targeting analysis in Hek-293H cells).

The general protocol followed is represented in **Figure 24**.

• Freezing/Thawing of HEK-293H cells in p96 plates and pseudoclones stablishment

In order to freeze the p96 plates, the media in the wells was aspirated and the cells washed with PBS (Dulbecco´s phosphate buffered saline, Sigma-Aldrich). After removing the PBS, 20 µL of Trypsine (Trypsin-EDTA solution, Sigma) were added and cells resuspended in 20 µL of Hyclone FBS with 40% DMSO (Dimethyl sulfoxide, Sigma-Aldrich) (final dilution of 20%) and kept at -80ºC.

The thawing was performed by adding 100-200 µL of DMEM + 10% HyClone + 1% P/S to each well and cells were resuspended. Then those cells were seeded in p12 plates in order to establish a pseudoclone as cell line. Cells were revised regularly and changed to p6, p100 or flask as the cells were proliferating. Once the culture is established, the cells were maintained in 75 flasks.

• Flow cytometry analyses: cell viability and $EGFP^+$ cells

Analysis was performed in the LSRFortessa cell analyser (BD/Becton, Dickinson and Company, New Jersey, USA). Off-line analysis was conducted with the FlowJo Software v7.6.5 (Tree star, Ashland, USA).

An aliquot of the pool of cells transfected with the nucleases or cells from the pseudoclones were washed and resuspended in the flow cytometry buffer (PBS containing 0.5% BSA and 0.05% azide) containing 3.3 ug/mL of 4',6-diamidino-2-phenylindole (DAPI) as a viability marker.

• gDNA extraction of the HEK-293H cells in p96 plates

The media from the wells was aspirated and 40 µL of a mix of Lysis buffer + Proteinase K were added per well and incubated at 55ºC 2 hours and at 95ºC 10 minutes. Then the plate was centrifuged and 3 µL of the supernatant were used to perform the PCR for the targeted integration analysis.

Lysis buffer composition for 500 mL: 495.5 mL of 10 mM Tris HCl pH 8 + 2.25 mL of 0.45% (v/v) NP40 (Nonidet P-40 substitute, Roche Diagnostics) + 2.25 mL of 0.45% (v/v) of Tween 20 (Sigma-Aldrich), storage at 4ºC. Per p96 plate, 5 mL of Lysis Buffer + 25 µL of Proteinase K (100 µg/mL final) were mixed in the day of the extraction

• Specific gene targeting analysis in HEK-293H

The PCR used to test the specific integrations of the GM or the TM mediated by nucleases in HEK-293H was performed using the Herculase II Fusion Enzyme (Agilent Technologies) and the primers showed in **Table 5**. This pair of primers amplifyed in the 3´side of the integration, the forward one binding inside the cassette and the reverse one in the genome, outside of the left-HA. Three µL of the gDNA were used for the PCR reaction following the vendors' guidelines, using the following program: 94ºC 10 minutes, 35 cycles of 94ºC 1 minute, 52ºC 30 seconds, 72ºC 2 minutes, and finally, final elongation at 72ºC 10 minutes. The PCR was tested in 1% agarose gels, resulting in a band of 1,565 bp when the donor is integrated in the SH6 target site, as shown in **Figure 35**.

Primer	Sequence 5´→ 3´
SH6 HR R-HA 1F	ACTCATCAATGTATCTTATC
SH6 HR R-HA 1R	TTGCAAAGTACATTGGTAAT

Table 5: Name and sequence of the primers used in the analysis performed to check the targeted integration of the donors in the SH6 locus.

In order to check the targeted integration of the GM in the SH6 locus in the established pseudoclones, an aliquot of the cells was harvested and the gDNA extracted using the NucleoSpin® Tissue kit. PCR analysis was performed using 150 ng of gDNA as described previously in this point

• Southern blot

The gDNA from samples of the established pseudoclones was obtained using the phenol-chloroform protocol (MANIATIS). 50 µg of genomic DNA were isolated from each studied pseudoclone and digested with *BamHI* (New England Biolabs, MA, USA) in separate reactions for 8 hours at 37°C. The digested DNA was concentrated by precipitation with ethanol and sodium acetate and was measured with Nanodrop®. 20 µg of each digestion were resolved on a 0.8% TBE agarose gel. The transfer to a NytranSuPerCharge nylon membrane was carried out overnight by turbo Blotter transfer system (Whatman, GE healthcare) following vendor guidelines and then incubated at 80ºC to fix the DNA to the membrane. Pre-hybridization and hybridization were performed by quick hybridization solution (Agilent technologies) in rolling cylinders inside the hybridization oven at 65°C. The probe used was the sequence of the EGFP gene, which was generated by digesting the GM Donor with *BamHI*/*SbfI* restriction enzymes and isolating the 892 bp band from an agarose gel. It was labeled with radioactive P^{32}dCTPs by Prime-It II Random Primer Labeling kit (Agilent) following vendor guidelines and the non-incorporated radioactivity was removed by cleaning the probe with IllustraNICK Columns (GE

Healthcare). After 15 minutes of prehibridization, the probe was added to the quick hybridization solution and left overnight. After 12 hours, the membrane was washed twice with 2X SSC, 0.1% SDS for 10 minutes at room temperature, then for 15 minutes with pre warmed 1X SSC, 0.1% SDS at 65°C and finally for 50 minutes with 0.1X SSC, 0.1% SDS at 65°C with two media changes. Then, the membrane was covered with plastic wrap and introduced in a Hyperscreen autoradiography cassette together with an Amersham Hyperfilm ECL (GE Healthcare) that was exposed overnight at -80°C covered. After a minimum of 10 hours, the film was developed in an automated processor.

4.2 Gene targeting in human CD34$^+$ cells

Human CD34$^+$ cells obtained from umbilical cord blood as described in the point 2.1 of Materials and Methods were thawed and pre-stimulated in pre-stimulation media during 48 hours at a density of 10^6 cells/mL. Then around 10^5 to 10^6 cells were nucleofected using the AMAXA Nucleofector kit for human CD34$^+$ cells or the P3 Primary Cell 4D-Nucleofector L or S Kit (Lonza), depending on the experiment. Different amounts of the MN or the TALEN (see RESULTS points 4 and 5) alone or with 4 µg of the GM donor were used in different experiments. Depending on the protocol performed, time in culture media varies and the percentage of EGFP$^+$ cells and primitive population were measured by Flow Cytometry. In some experiments cells were sorted later. Methylcellulose assays were performed; colonies were picked and checked for specific integration of the GM donor in the SH6 donor by NESTED PCR.

All these steps were performed as explained below.

- **Flow cytometry analysis: checking cell viability, EGFP$^+$ cells, primitive populations, cell cycle and sorting of CD34$^+$/EGFP$^+$ cells**

Cytometry analysis was performed in the LSRFortessa cell analyser (BD/Becton, Dickinson and Company, New Jersey, USA). Off-line analysis was conducted with the FlowJo Software v7.6.5 (Tree star, Ashland, USA). The sorting was conducted in the Influx BD ®.

For the primitive population analysis, 2*10^5 CD34$^+$ nucleofected cells with the nuclease and donor combination were harvested at different time points and resuspended in flow cytometry buffer. After cells were resuspended in 100 µL of flow cytometer buffer and stained during 30 minutes at 4°C with CD34 PE-Cy5, CD38-PE, CD45RA-APC eFluor 780 and CD90-APC. Then, cells were washed and resuspended again in the flow cytometry buffer containing 3.3 ug/mL 4',6-diamidino-2-phenylindole (DAPI) as a viability marker (**Table 6**).

Antibody	Clone	Host	Supplier
CD34-PECy5	581	Mouse IgG1, κ	Immunotech
CD38-PE	H37	-	BD
CD45RA-APC eFluor 780	HI100	Mouse IgG2b, κ	eBioscience
CD90-APC	5E10	Mouse IgG1, κ	BD
CD34-PE	8G12	-	BD

Table 6: List of the antibodies used in the FACS analysis of the primitive populations in nucleofected $CD34^+$ cells.

In this analysis the percentage of $EGFP^+$ cells was measured in the total population of cells and also in the different subpopulations marked with different antibodies. Gating strategy is shown in **Figure 48**: $CD34^+/CD38^-$ cells were selected at first and, inside this population, $CD45RA^-$ cells were gated. The most primitive compartment, the HSCs (long term engraftment) correspond to the positive cells to CD90 ($CD34^+/CD38^-/CD45RA^-/CD90^+$). On the contrary, the $CD90^-$ cells correspond to MPP (short term engraftment) ($CD34^+/CD38^-/CD45RA^-/CD90^-$) (Doulatov, Notta et al. 2012).

In the sorting experiments, cells (not used in the pre-sorting clonogenic assays) were resuspended in flow cytometry buffer and stained during 30 minutes at 4°C with CD34 PE (**Table 6**) to isolate the $CD34^+/EGFP^+$ cells.

Cell cycle analysis was performed by washing the cells in PBS + BSA 10% and resuspending them in 200 µL of cold EtOH 70% to be fixed for at least 40 minutes. After washing in PBS, the cells were resuspended in 200 µL of PBS with 50 µg/mL of RNAsa during 20 minutes at 4ºC. Then, 50 µg/mL of PI (Propidium Iodide, Sigma) is added to the mix and incubated 15 minutes at RT in darkness before the analysis by Flow Cytometry.

• Clonogenic assays

Depending on the experiment, 300 to 1,500 $CD34^+$ cells were resuspended in 3 mL of methylcellulose. Each mL of the triplicate was seeded in a M35 plate and cells were incubated in normoxia: 37°C, 21% O_2, 5% CO_2 and 95% RH. Fourteen days later, the number of colonies was counted using an inverted microscope (Olympus IX70 WH10x/22, objective 4x) and the CFU-GMs (granulocyte-macrophage colony forming unit) and BFU-Es (Burst Forming Unit-Erythroid) were identified. Single colonies were picked to analyze by specific PCR the integration of the GM donor in the *SH6* locus.

- ## gDNA extraction of the colonies obtained from clonogenic assays cells in p96 plates

Picked single colonies were distributed in p96 plates and washed in 100-200 µl of PBS, 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) and incubated at 65°C for 30 min, 90°C for 10 min and 4°C. After the lysis, 30 µl of water was added and the mixes were centrifuged.

- ## Specific gene targeting analysis in human $CD34^+$ cells: NESTED-PCR

Nested-PCR was used to demonstrate that $EGFP^+$ colonies carried a specific integration in SH6 locus using the Herculase II Fusion Enzyme (Agilent Technologies) PCR enzyme. The Nested-PCR consists in two PCRs: In the Nested1, we used a pair of primers targeting inside the GM and in the genome far away from the right HA (R-HA) **Figure 22**. In the Nested2, the product of the Nested1-PCR was used as template and the pair of primers was the same as we used in the PCR in experiments in HEK-293H cells **Figure 22**, and targeted inside the Nested1 amplicon. The primers used are summarized in **Table 7**.

***SH6* locus with the integrated GM**

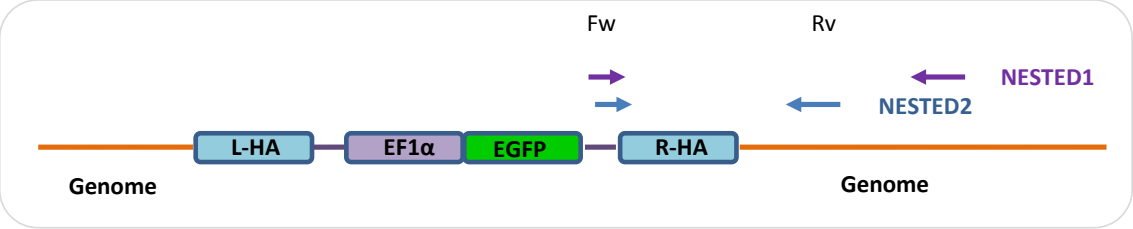

Figure 22: The scheme represents the Nested PCR strategy. Purple arrows represent the pair of primers corresponding to the first PCR (Nested1) and the blue ones, those corresponding to the second one.

Ten µL of the gDNA obtained from the CFC/Hematopoietic progenitors, was used to perform the NESTED1 PCR with the PCR program: 94ºC 10 minutes, 35 cycles of 94ºC 1 minute, 66ºC 30 seconds and 72ºC 2 minutes, and finally one cycle of final elongation at 68ºC 10 minutes. Then, 2 µL of the resulted PCR product was used as template for the NESTED2 using the same program as with the HEK-293H cells: 94ºC 10 minutes, 35 cycles at 94ºC 1 minute, 52º 30 seconds, 72ºC 2 minutes and finally, one cycle at 72ºC 10 minutes. The result was resolved in a 1% TBE agarose gel.

Primer	Sequence
NESTED PCR HR R-HA GM 1F	TTCGGATCCGATAATATGGCCACAA
NESTED PCR HR R-HA GM 2R	TGCAGACATTAGATTAGGACTTGGGTTA
SH6 HR R-HA 1F	ACTCATCAATGTATCTTATC
SH6 HR R-HA 1R	TTGCAAAGTACATTGGTAAT

Table 7: Name and sequence of the primers used in the NESTED-PCR analysis performed to check the targeted integration of the donors in the SH6 locus.

In order to check the targeted integration of the GM in the SH6 locus in the pool of cells, an aliquot of the nucleofected cells was also harvested and the gDNA extracted using the NucleoSpin® Tissue kit. PCR analysis was performed using 150 µg of gDNA as described previously in this point.

4.3 Nucleases functionality or cleavage assay (Surveyor assay)

The samples of HEK-293H or $CD34^+$ cells transfected or nucleofected with the nucleases as described in Materials and Methods points 4.1 and 4.2 were harvested and the gDNA extracted using the NucleoSpin® Tissue kit. The aim of this assay is to detect the possible mismatches originated by the repair of the double strand breaks (DSB) generated by the cut of the nucleases in the target site. When this DSB is repaired by NHEJ, mutations as insertions and deletions are generated, called INDELs. At first, PCR was performed using a pair of primers (**Table 8**) binding in the middle of the HAs in order to amplify the target region, obtaining a 351 bp band and using Herculase II Fusion Enzyme. The program employed consisted in one cycle at 95ºC 5 minutes, followed by 33 cycles at 95ºC 20 seconds, 60ºC 25 seconds and 68ºC 30 seconds, and finally one cycle of final elongation at 68ºC 10 minutes. This band was denaturalized and rehybridized in order to form heteroduplex between mutated and non-mutated strands. If an INDEL has occurred, the heteroduplex presents a DNA loop that can be recognized by the Surveyor® nuclease (Surveyor® mutation detection kit, IDT, Coralville, Iowa, USA) CelI and cut, generating a band pattern that was visualized on 10% TBE Gels 1.0 mm (Invitrogen). A positive result is represented by the appearance of two additional bands of around 179 and 171 bp (example in **Figure 29B**). The intensity of these bands was measured using the Quantity One program from Bio-Rad, compared with the intensity of the first one of 351 bp in the top. This intensity was indicative of the percentage of activity of the nuclease. The percentage of cleavage is determined by the next equation:

$$\%\ NHEJ = \frac{Cleavaged\ bands - (2 \times Background)}{(Cleavaged\ bands + Parental\ band) - (3 \times Background)} \times 100$$

Primer	Sequence
SH6 SurvAssay 1F	TCTTTGTGTTTCCAAAGAGTTCCTTTGGCTTTC
SH6 SurvAssay 1R	GAATGGTCTGAAAATGGAGAGGTTAAATGAGATTT

Table 8: Name and sequence of the primers used in the Surveyor assay performed to check the activity of the nucleases.

RESULTS

1. COMPARISION OF THE EFFICIENCY OF DIFFERENT MEGANUCLEASES TARGETING THE SH6 SAFE HARBOR IN HEK-293H CELLS

Two specific meganucleases (MNs) were generated by Cellectis targeting safe harbor known as the Safe Harbor 6 (SH6). This SH6 was found by Cellectis in a genomic desert located in chromosome 21 and may constitute an alternative to other SH nowadays used in targeted gene therapy. These MNS are named SH6-meganucleases (SH6-MN). Two of them, SH6v2 and SH6v5, have been tested in our studies in the HEK-293H cell line to identify which one was less toxic, safer and had a better efficiency to induce homologous recombination (HR) events in this target site. A donor DNA carrying the EGFP reporter gene, here referred as the Green Matrix (GM), was used together with the nucleases to evaluate HR events in the target site.

1.1 The Green Matrix is highly expressed in HEK-293H cells

The first donor plasmid that was used in our HR experiments was the GM. This plasmid carries an EGFP reporter gene under the regulation of the long EF1α promoter. Both sequences were flanked by homology arms (HA) to the SH6 locus. The EGFP reporter gene would allow us to identify those cells which had integrated of the GM.

First we analyzed the functionality of the GM by measuring the expression of the transgene. Cells were transfected either with 4 µg of a control plasmid carrying both, the EGFP and the Cherry reporter genes (EGFP/Cherry plasmid) under the PGK promoter, or with 4 µg the GM, using Lipofectamine 2000 **Figure 23A**.

The expression of EGFP was measured by FACS at days 3, 11, 20 and 24 post-transfection in HEK-293H cells. On days 3 and 11 the percentage of $EGFP^+$ cells was always higher in GM transfected cells compared with the EGFP/Cherry transfected cells (97% vs 67% on day 3 and 37% vs 1.7% on day 11, respectively). Also a higher intensity of expression was observed in the GM transfected cells, due to the probably higher number of copies of the GM, compared to the GFP/Cherry, transfected. The percentage of $EGFP^+$ cells decreased as the days passed, indeed due to the dilution of the plasmid. By day 20 post-transfection, almost no $EGFP^+$ cells were observed, regardless of the plasmid transfected (**Figure 23B**).

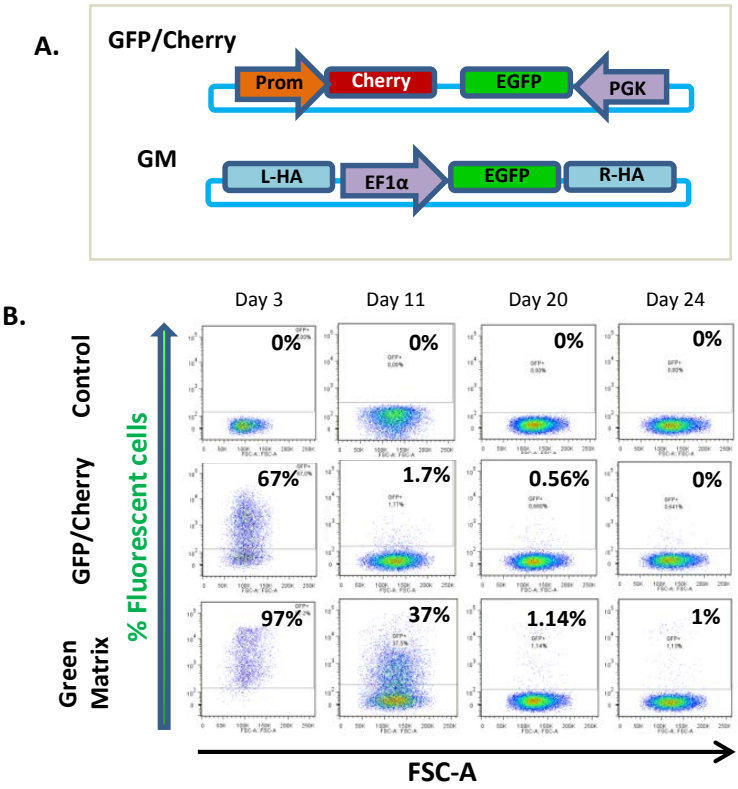

Figure 23: **A.** GFP/Cherry and GM plasmids schemes are represented. The plasmid GFP/Cherry carries Cherry and GFP genes. The GM carries an EGFP reporter gene under the long EF1α promoter. The two components are flanked by 1.5 kb sequences with homology to the *SH6* locus. **B.** FACS analysis representing the EGFP+ cells at different time points after transfection.

Summarizing these results, higher percentages of $EGFP^+$ cells expression are observed when the HEK-293H cells were transfected with the GM, and both plasmids were diluted when 15 days passed.

1.2 Induction of Homologous Recombination of the GM using SH6-meganucleases

In our next experiments we compared the efficiency and specificity of either, SH6v2 and SH6v5 meganucleases to induce HR events of the GM. 0.5 µg of plasmid DNA of either SH6v2 or SH6v5, were transfected in HEK-293H cells, with 4 µg of the GM. At 3 days post-transfection, cells were counted and 10 cells per well were plated in p96 plates. The rest of the cells were maintained in culture to analyze the percentage of EGFP+ cells by FACS up to 37 days post-transfection. In parallel, cells in were cultured in p96 plates for 7-15 days, and then duplicated. After 1-2 weeks, half of the plates were frozen. DNA was extracted from the other plates to quantify the HR events by PCR using specific primers binding sequences inside and outside of the cassette. After identifying wells carrying cells with the specific integration, the cells from the frozen plates were used to stablish pseudoclonal cell lines. From these cells, PCR, FACS and Southern blot analyses were performed. This protocol is explained in **Figure 24.**

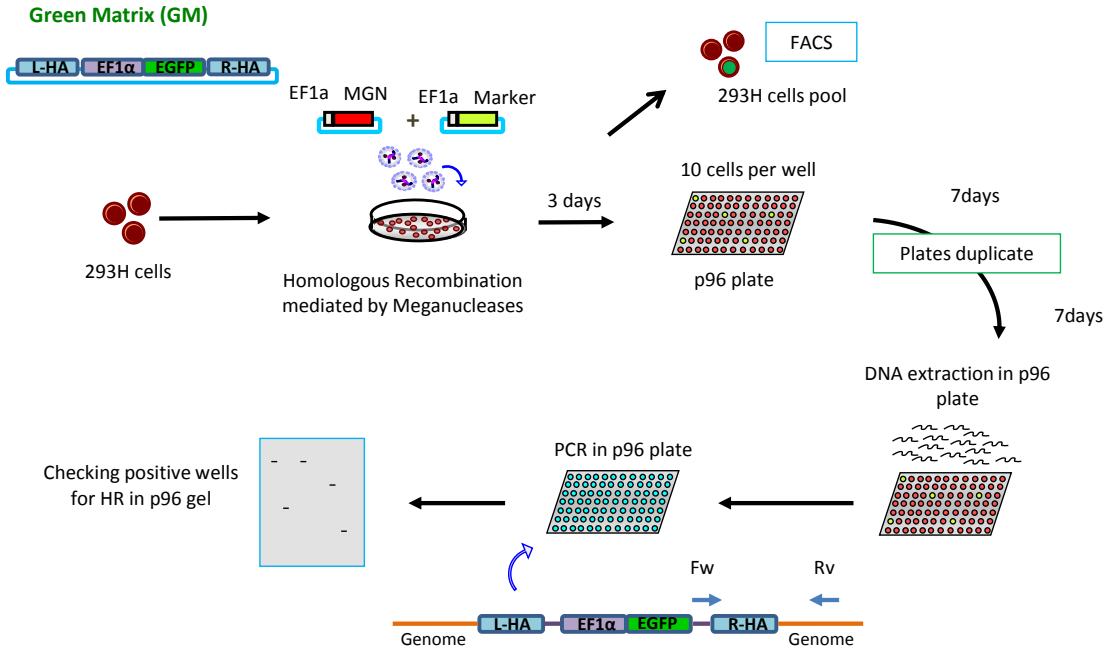

Figure 24: Schematic representation of the experimental protocol used to induce homologous recombination of the GM in the SH6 site, mediated by SH6-MNs, in HEK-293H cells.

- **Sustained EGFP expression in cells co-transfected with the SH6v5 meganuclease and the GM**

The percentage of EGFP+ cells over time was investigated in samples transfected with the MNs and the GM or with the GM alone. The percentage of $EGFP^+$ cells was always until day 11. At day 24 only the 0.5 µg SH6v5 condition presented a 4% of $EGFP^+$ cells, suggesting the integration of the EGFP gene in the *SH6* locus, and maybe by HR. When the SH6v2 was used, no difference in the percentages of $EGFP^+$ cells was observed with comparing to the Control and GM conditions **Figure 25.**

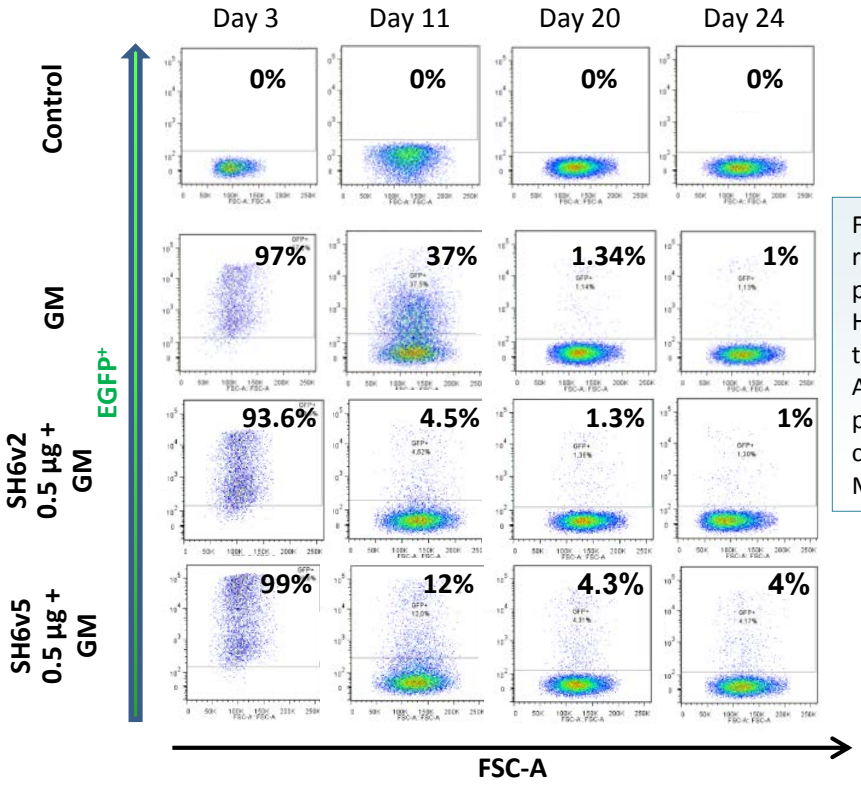

Figure 25: Dot blots representation of the percentages of EGFP$^+$ HEK-293H cells transfected over time. At day 24, higher population corresponds to SH6v5 MN.

Cell transfected with the nucleases and the GM and after more than 15 days in culture maintained a higher percentage of EGFP+ cells. This could be indicative of the EGFP transgene integration in the genome, and maybe by HR.

- **Study of the efficiency of Homologous Recombination in the SH6 site using SH6-meganucleases and the GM**

In order to confirm that HR occurred at the SH6 site, HEK-293H cells were plated in 96 wells-plates (10 cells by well) three days after transfection. 2-4 weeks later, DNA from these cells was extracted. In order to check the percentage of HR events, PCR performed using a pair of primers targeting inside and outside the cassette. A positive result is indicative of specific integration in the SH6 site (**Figure 26A)**.

Similar percentages of HR events, 9.89% and 6.4%, were obtained using both meganucleases, SH6v2 and SH6v5, respectively, with the dose of 0.5 µg. These percentages are obtained by scoring the number of positive wells for the specific PCR among the 96 wells of the plate. The clonogenic capacity of these cells must be taken into account, because it is estimated that 3 cells from the 10 plated are growing (DATA NOT SHOWN), so these percentages of HR could be up to 3 times lower.

This result indicated that both versions of SH6-MNs could induce HR in the SH6 site in 293H cell line. **Figure 26B.**

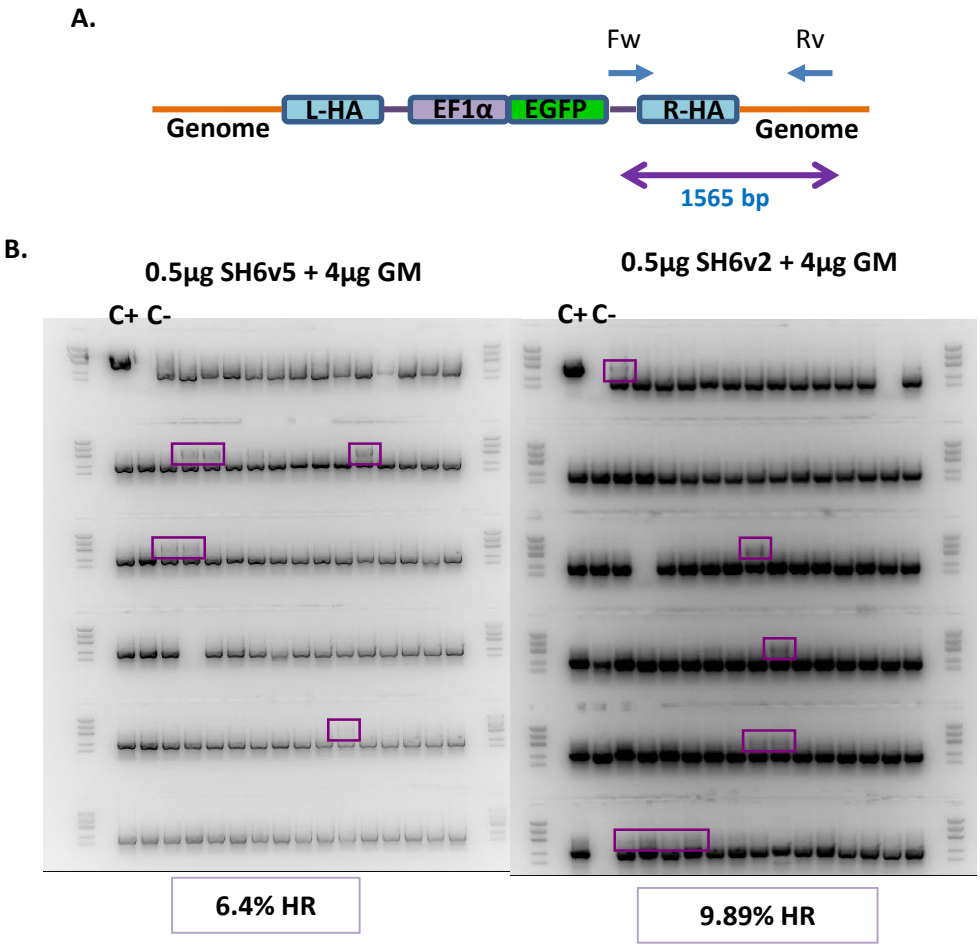

Figure 26: **A.** The scheme represents the GM integrated in the SH6 locus. Blue arrows represent the pair of primers used in the PCR analysis to test the specific integration of the GM in the target site. The purple arrow shows the size of the amplicon obtained as a positive result in the PCR. **B.** Agarose electrophoresis gels showing positive bands, inside the purple boxes, corresponding to HR events in p96 plates.

- **SH6v5-MN induces specific and stable integration of the GM in the SH6 site in HEK-293H cells**

Once pseudoclones containing the GM integrated in the target site were established from the wells positive for HR events, further analyses were performed in order to confirm the specific integration of the GM. These pseudoclones are a mixture of cells carrying HR events and cells without integrations.

FACS analysis showed the presence of $EGFP^{+}$ cells in the established pseudoclones, indicating that long EF1α promoter was able to drive stable EGFP expression **Figure 27A**. Differences in the intensity and the percentage of EGFP+ cells in each pseudoclone indicated different expression profile in targeted cells.

PCRs for the specific integration were also performed in cultured pseudoclones. The presence of specific bands indicated that the integration was specific, stable and not toxic for the cells. **Figure 27B**.

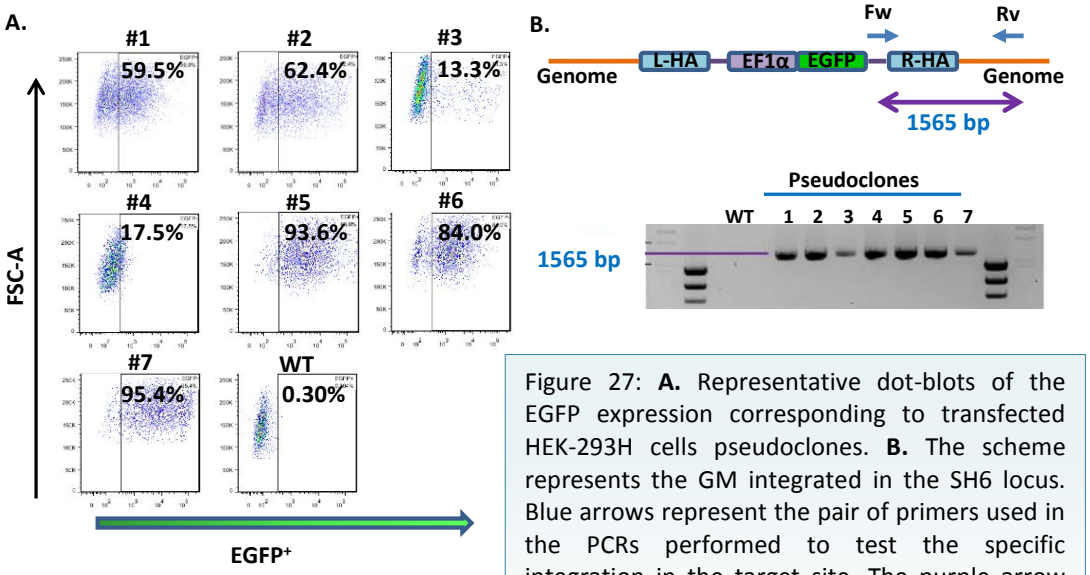

Figure 27: **A.** Representative dot-blots of the EGFP expression corresponding to transfected HEK-293H cells pseudoclones. **B.** The scheme represents the GM integrated in the SH6 locus. Blue arrows represent the pair of primers used in the PCRs performed to test the specific integration in the target site. The purple arrow indicates the size of the amplicon corresponding to the positive band in the PCR. The agarose gel electrophoresis shows the positive band in tested pseudoclones. These analysis were performed almost 6 weeks post-transfection.

Established pseudoclones obtained from transfected cells and positive for HR events showed different intensities and percentages of EGFP+ cells, consistent with the heterogeneous nature of the pseudoclones.

1.3 The SH6-MNs did not induce off-target integrations of the GM in HEK-293H cells

In order to determine if SH6-meganucleases induced off-target integrations of the GM, pseudoclones generated after transfection and positive for HR with the GM were analyzed by Southern blot.

The sequence of the EGFP transgene employed in the GM was labeled with P-32, and used as a probe. DNA from pseudoclones was digested with BamHI, with a target site inside the cassette and another one outside of the HAs (expected band size of 3,946 bp **Figure 28A**.

The expected band of 3,946 bp was generated in pseudoclones 1, 2, 4, 5, 6 and 7 after cleavage using BamHI, indicating the specific integration of the GM in the SH6 site, An additional band of around 7,242 bp was observed in pseudoclones 1, 5, 6 and 7 (**Figure 28B)**.

This band could correspond to additional way integrations of the GM due to the use of a circularized donor plasmid (see Discussion). **Figure 28C**.

When samples were digested with ScaI, the expected 2,738 bp band was observed in all pseudoclones, but not in pseudoclone 4, in which a smaller band was observed.

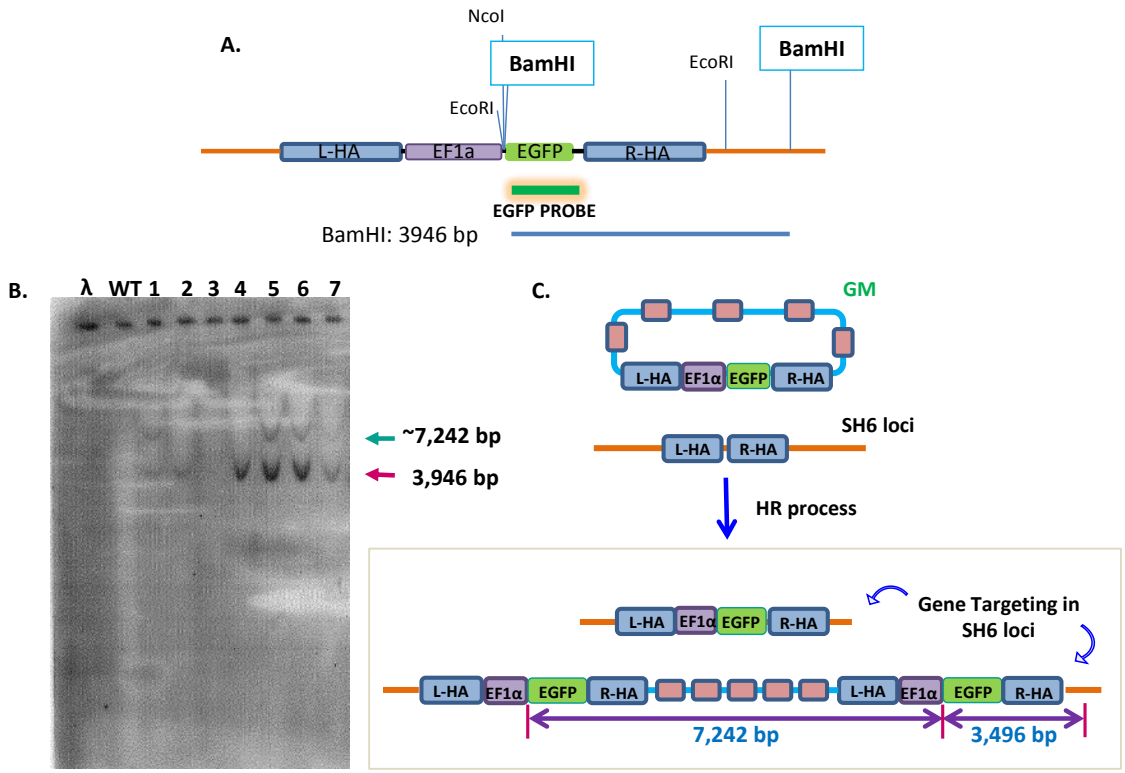

Figure 28: **A.** Southern blot strategy: Hybridization with a radioactive labeled probe that binds the EGFP gene revealed a band of 3,946 bp after digestion of the genomic DNA (gDNA) with BamHI. **B.** Southern blot film corresponding to digested gDNA of pseudoclones 1 to 7. The red and green arrows mark, respectively, the expected size bands and the additional bigger bands explained in **C. C.** Representative scheme of the GM, a circular plasmid donor, and the possibilities of HR mediated integration in the target site. The scheme inside the square represents, in the top, the desired integration with the cassette alone, and below, another possibility that could explain the bigger and repetitive band among the pseudoclones.

2. COMPARISION BETWEEN A SH6-TALEN AND THE SH6v5 MEGANUCLEASE TO MEDIATE TARGETED INTEGRATION OF THE GM IN HEK-293H CELLS

In the following studies, toxicity and efficiency for inducing HR events of another nuclease, a TALEN designed to target the SH6 (SH6-TALEN), was compared to the SH6v5-MN performing analyses following the same protocol used before and detailed in **Figure 24.**

2.1 Analysis of the Homologous recombination events induced by the SH6-TALEN in HEK-293H using the GM

SH6-TALEN was transfected into HEK-293H cells using two plasmids, each one codifying for one of its subunits. Four doses of each plasmid (0.3, 0.5, 3.0 and 10.0 µg), plus 4 µg of the GM, were used and its effects compared with those mediated by 0.5 µg of the SH6v5-MN plus 4 µg of the GM.

- **SH6v5-MN and SH6-TALEN nucleases functionality comparition in the *SH6* locus**

First step to compare both nucleases, the SH6v5-MN and the SH6-TALEN, was to test their activity in the cells targeting *SH6* locus. HEK-293H cells were transfected as described and 3 days later DNA was extracted to perform the surveyor assay. The aim of this assay is to detect mismatches originated by the repair of double strand breaks (DSB) generated by the nucleases in the target site. At first, PCR was performed using a pair of primers binding in the middle of the HAs in order to amplify the target region, represented by a 351 bp band. This band was denaturalized and rehybridized to form heteroduplex between mutated and non-mutated strands. This heteroduplex was digested with the CelI enzyme, and the product was visualized by electrophoresis in TBE agarose gels. A positive result is represented by the appearance of two additional bands of around 179 and 171 bp. The relative intensity of these bands, compared with the band of 351 bp in the top, was indicative of the percentage of activity of specific cleavage of the nuclease in the SH6 site.

As shown in **Figure 29A**, bands present different intensities depending of the nuclease and the dose. The measure of the relative intensity of the different bands in **Figure 29C** represents the percentages of specific cleavage obtained in each condition. These results highlighted that all

the nucleases showed similar activity and were able to induce mutagenesis in the SH6 target site.

- **Maintained expression of EGFP gene in the pool of cells transfected with the SH6v5 meganuclease or the SH6-TALEN and the GM**

Analyses by FACS were performed in order to evaluate the percentage of EGFP+ cells in the pool of transfected cells. At day 37 post-transfection, the EGFP percentage was around 2-2.5% in cells transfected with 0.5, 3.0 or 10 µg of TALEN plus the GM. A 2 fold of increase of $EGFP^+$ cells was observed in cells treated with the higher dose of the TALEN and the MN compared with the cells transfected with the GM alone, suggesting that the cleavage of nucleases could facilitate the specific integration of the EGFP gene in the SH6 site (**Figure 29C**).

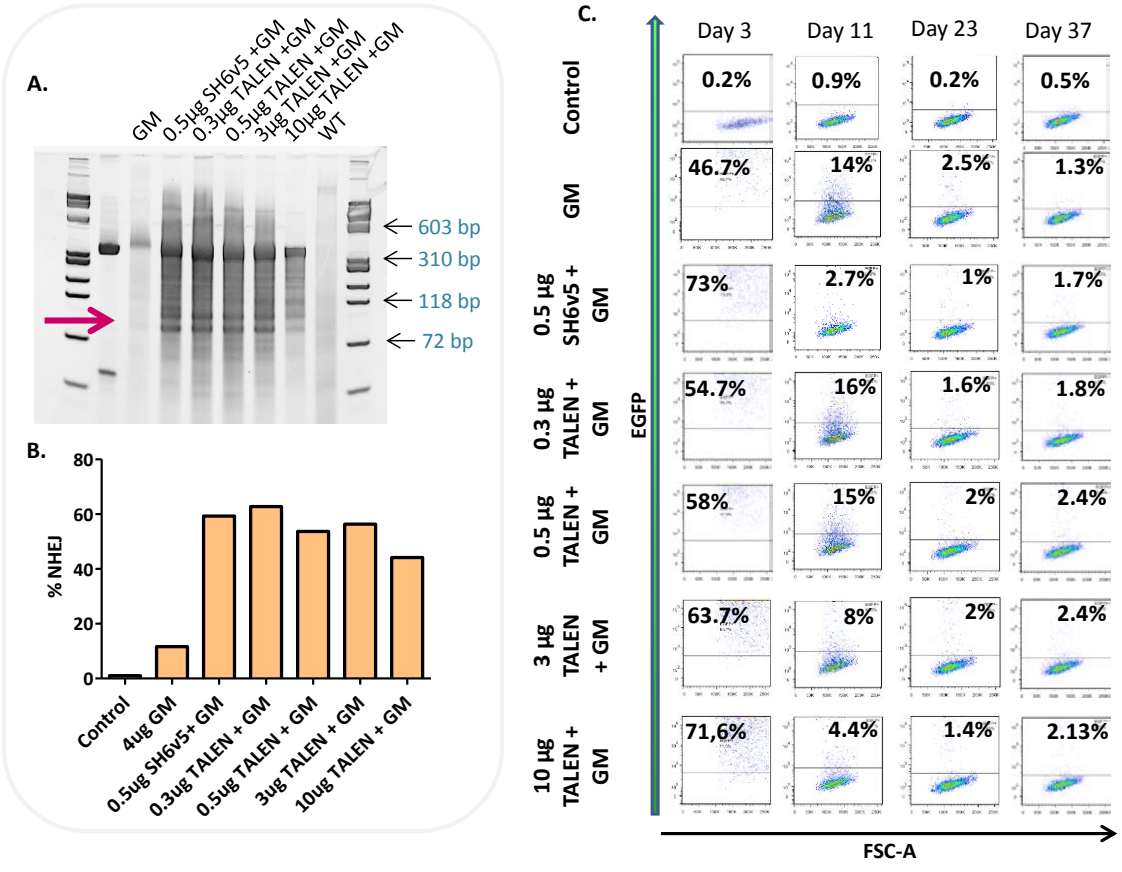

Figure 29: **A.** Surveyor analysis in the pool of cells digested with the enzyme CelI visualized by electrophoresis. The red arrow marks the two bands are indicative of the activity of the nucleases. **B.** Quantification of the intensity of the bands showed in **A.**, indicating the percentage of specific cleavage. **C.** Representative panel of the percentage of $EGFP^+$ in HEK-293H cells transfected with the GM alone or with the nucleases along time.

- **Comparative analysis of the targeted integration of the GM mediated by the SH6-TALEN and the SH6v5-MN**

In order to analyze the percentage of HR events, HEK-293H cells transfected with the SH6-TALEN and the SH6v5-MN were plated in 96 well-plates (10 cells by well), and DNA was extracted 3 days after transfection to perform PCR analyses with primers targeting inside and outside of the GM to determine the proportion of targeted integration.

As shown in **Figure 30**. Efficiencies of HR in the range of 15-20% were achieved with 0.5 µg of the SH6v5-MN and with 3 or 10 µg of the SH6-TALEN.

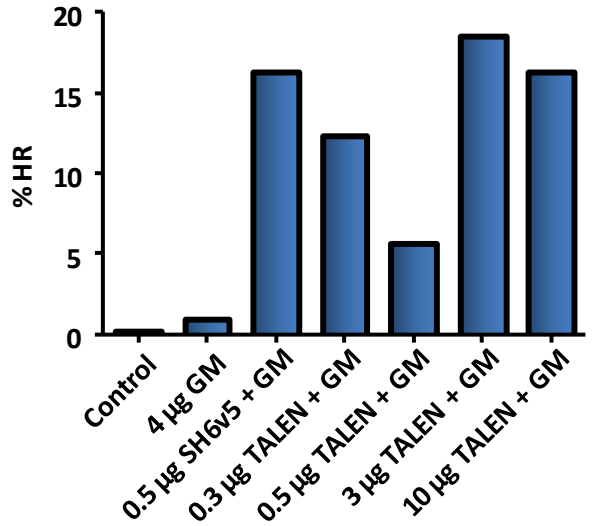

Figure 30: Percentages of HR obtained after co-transfection of the GM either with the SH6-MN and different doses of the SH6-TALEN in HEK-293H cells.

- **Both SH6v5-MN and SH6-TALEN mediate the specific and stable integration of the GM in the SH6 site in HEK-293H cells**

As in the experiments performed with the MNs, pseudoclones were established with transfected cells that showed HR in specific wells. New analyses were performed to confirm if the HR events were specific for the SH6 site and maintained along the time in culture in these cells.

Cytometry analysis showed that not all the pseudoclones maintained the expression of EGFP (**Figure 31A)**. Analyses aiming to detect the specific integration of the GM were performed. Positive results were obtained only in clones with a high percentage of EGFP+ cells (**Figure 31B)**, indicating that the EGFP reporter gene is only expressed when it is integrated in the target site. In these positive pseudoclones the expression was stable and not toxic for the cells.

Nevertheless, the percentage of EGFP$^+$ cells was different in each pseudoclone consistent with their heterogeneous nature.

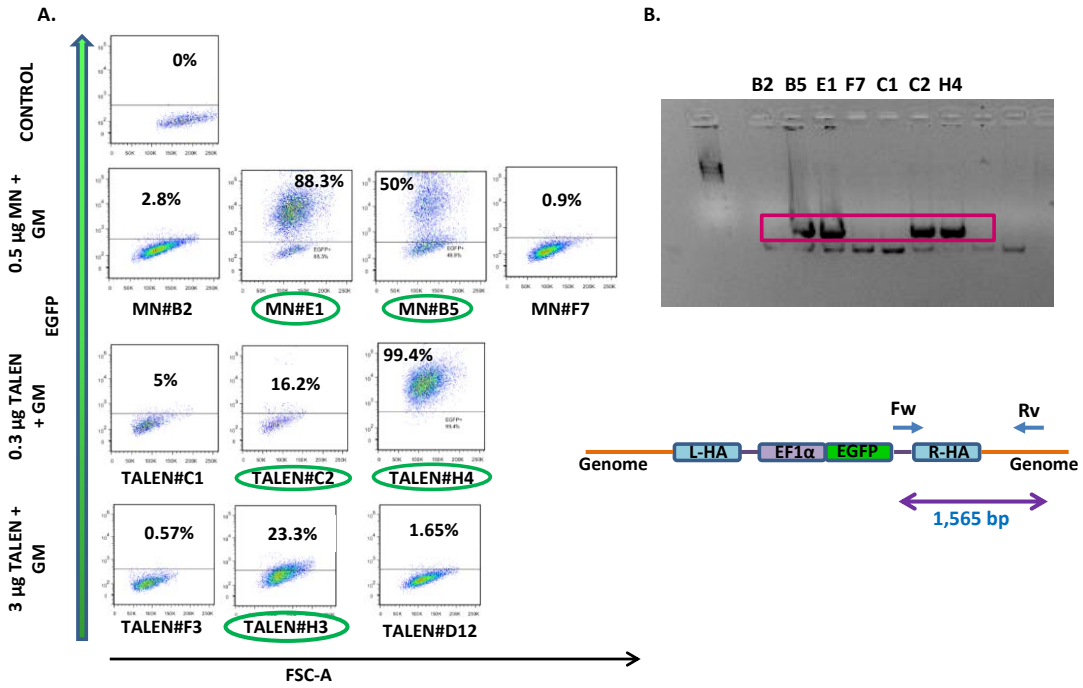

Figure 31: **A.** The panel shows the presence of EGFP$^+$ cells in some of the stablished pseudoclones. Pseudoclones positive for the specific integration of the GM are rounded in green. In **B.** the agarose gel electrophoresis represents which of the pseudoclones carry a specific integration of the GM, represented by a band of 1,565 bp as the purple arrow in the scheme below.

2.2 The SH6-TALEN does not induce off-target integrations

Southern blot analyses were performed to investigate the presence of targeted integrations and potential off-targets in HEK-293H cells transfected with the SH6-TALEN (**Figure 32**).

DNA from pseudoclones was digested with the BamHI restriction enzyme, with target sites inside and outside of the cassette (**Figure 32A)**. After hybridization with the labeled probe, the expected band of 3,946 bp was observed in the pseudoclones 2, 3 and 7, indicating the specific integration in the SH6 site. An additional band of 7,242 bp was observed in these clones (**Figure 32B**). This band could be the result of a different integration pattern of the cassette in the target site as result of the use of the circularized donor plasmid (**Figure 32C**). Surprisingly, no band was obtained in pseudoclone 6, in which an specific integration by PCR was identified after the culture of this pseudoclon. In this case, however, the level of EGFP expression was very low. No band was detected in pseudoclones 1, 4 and 5, consistent with the absence of the band in the previous PCR analyses, and with a low expression of EGFP (**Figure 31**).

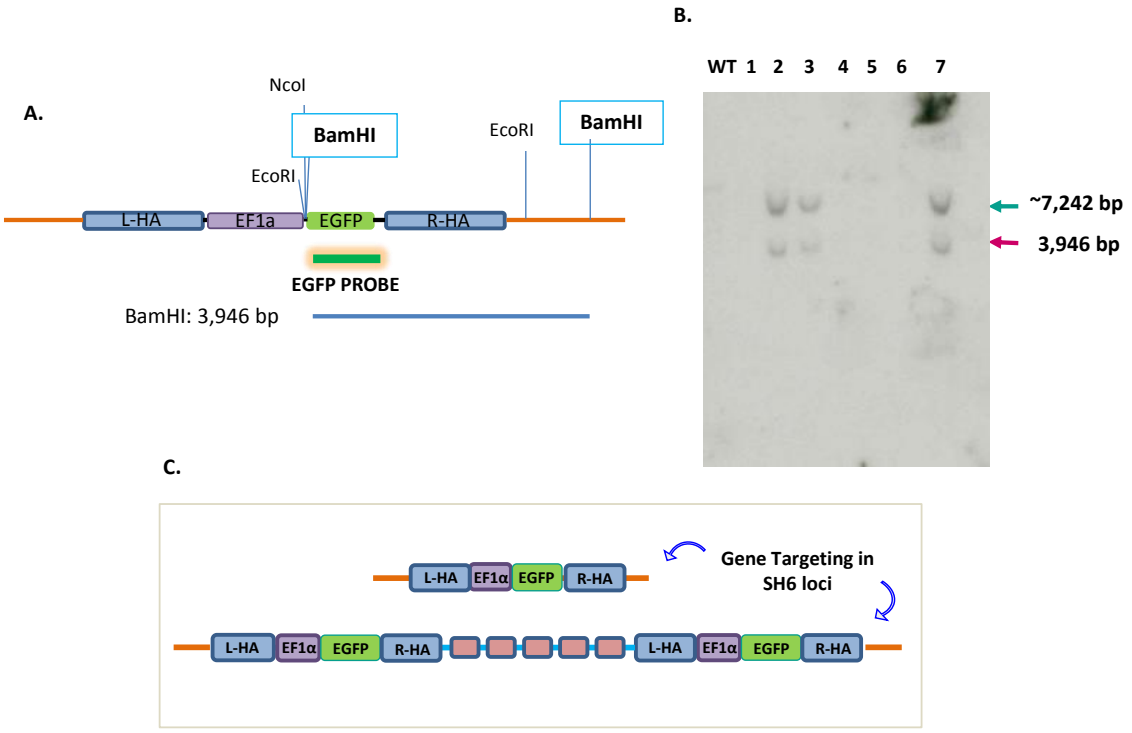

Figure 32: **A.** Southern blot strategy: a radioactive labeled probe that binds the EGFP gene generates a band of 3946 bp, after digestion of the genomic DNA (gDNA) with BamHI, if the targeting process takes place at the *SH6* locus**. B.** Southern blot film corresponding to digested gDNA of some pseudoclones. The red and green arrows mark, respectively, the expected size bands and the additional bigger band explained in Figure 6. **C.** In the bottom, scheme representing of the possibilities of HR mediated integration in the target site.

3. TARGETED INTEGRATION OF *FANCA* IN THE SH6 SITE IN HEK-293H CELLS USING SH6-TALEN AND SH6v5-MN

Once we had optimized the targeted integration of a reporter GM in the SH6 site of the HEK-293H cells, we aimed to develop a similar approach with a donor harboring the *FANCA* gene that we called the Therapeutic Matrix (TM). In this matrix the *FANCA* gene is driven by the EF1α promoter and followed by an E2A sequence and a puromycin resistant gene. As in the GM this construct is flanked by the two HAs for the SH6 target site (**Figure 33**).

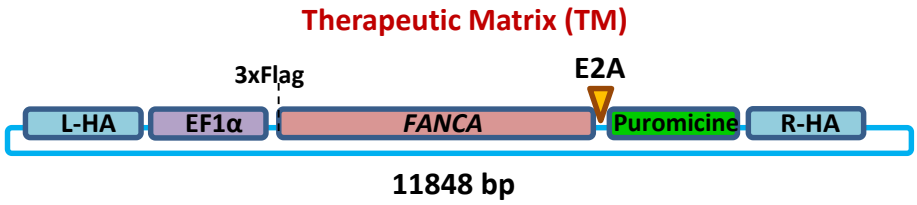

Figure 33: This scheme represents the Therapeutic matrix (TM) which carries the FANCA gene driven by the EF1α promoter and followed by an E2A sequence and a puromycin resistant gene. This construct is flanked by the two HAs for the SH6 target site.

3.1 HEK-293H cell viability after transfection with the Green and Therapeutic Matrixes and the MN and the TALEN nucleases

In order to investigate if the transfection with the TM (size 11,848) was affecting the viability of the cells, compared to the GM (size 7,440), samples were transfected with 4 µg of the GM or the TM alone or with 0.5 µg of the SH6v5-MN or 3 µg of the SH6-TALEN. Three days after transfection, the cell viability was compared among all conditions. As shown in **Figure 34**, transfection with the TM indicated that this matrix was not really cytotoxic for the transfected cells.

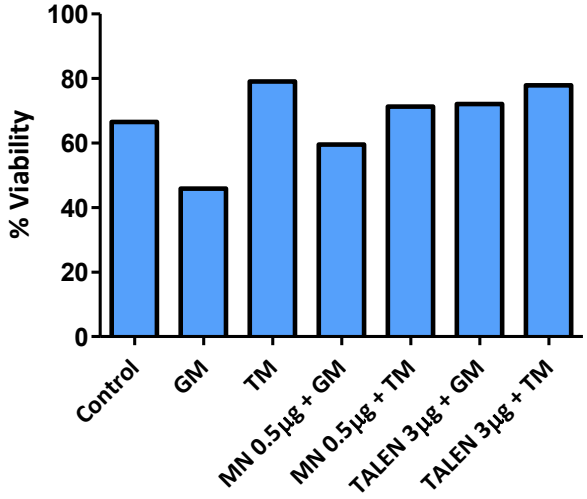

Figure 34: The figure represents the viability corresponding to cells transfected with the two matrixes and the nucleases, at three days post-transfection. In all instances, the viability is represented by the proportion of DAPI-negative cells.

3.2 MN and TALEN mediated targeted integration of the TM in the *SH6* locus in HEK-293H cells

In a next set of experiments we investigated the feasibility of performing targeted integrations of the TM in the SH6 site in HEK-293H cells. Transfected HEK-293H cells were plated in 96 wells-plates, (10 cells per well) and several analyzed according to the protocol represented in **Figure 24**.

When the GM was used, HR events were obtained with both nucleases, being the percentage of HR of 4% for the SH6v5-MN, and 1% for the SH6-TALEN. When the TM donor was used, HR percentages decreased to 1.8% for the MN and to <0.3% for the TALEN **Figure 35**. These results indicated that the use of a larger donor decreased the HR events. Nevertheless, were able to induce the specific integration of the TM when the SH&v5-MN was used.

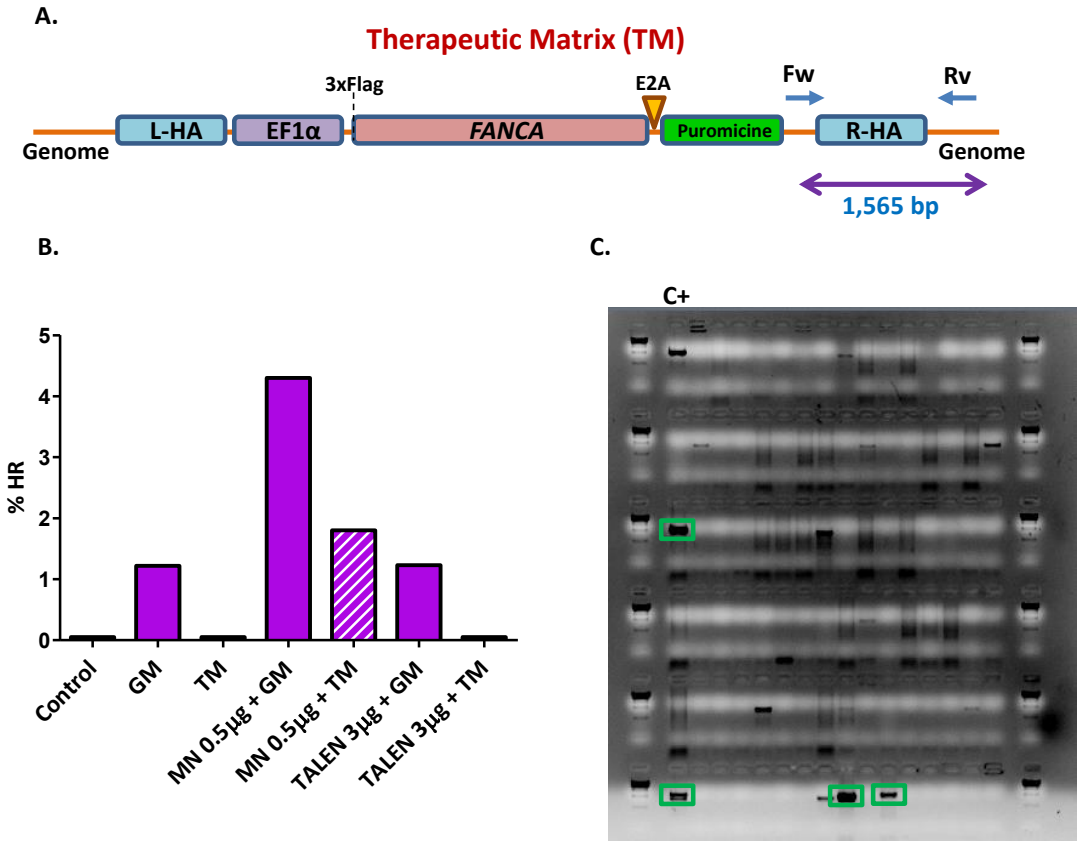

Figure 35: **A.** The scheme represents the TM integrated in the SH6 locus. Blue arrows represent the pair of primers used in the PCRs performed to test the specific integration in the target site. The purple arrow indicates the size of the amplicon corresponding to the positive band in the PCR. **B.** HR percentages induced by the MN and the TALEN using the GM or the TM donors. The lined bar represents the condition transfected with TM with positive HR. **C.** Agarose electrophoresis gels showing positive bands, inside the green boxes, corresponding to HR events in the MN+GM condition.

4. HOMOLOGOUS RECOMBINATION EXPERIMENTS USING SH6-MNs AND TALEN IN PRIMARY HUMAN CD34$^+$ HEMATOPOIETIC PROGENITORS

4.1 Analysis of Homologous Recombination mediated by nucleases in CD34$^+$ hematopoietic progenitor cells subjected to cell sorting of EGFP$^+$ cells

In order to investigate the feasibility of performing targeted gene therapy in the *SH6* locus of human HSCs, we designed the experimental protocol shown in **Figure 36**: Thawed cord blood (CB) CD34$^+$ cells were pre-stimulated for 48 hours and then nucleofected with the MN or the TALEN plus the GM. Seven days after nucleofection, the percentage of EGFP$^+$ cells was analyzed by FACS to determine the nucleofection efficiency. In order to enrich the targeted cells, EGFP+ cells were sorted. Methylcellulose assays were performed with the sorted cells and after 14 days, green colonies were picked. DNA from these colonies was extracted and PCR analyses were used to detect specific integrations.

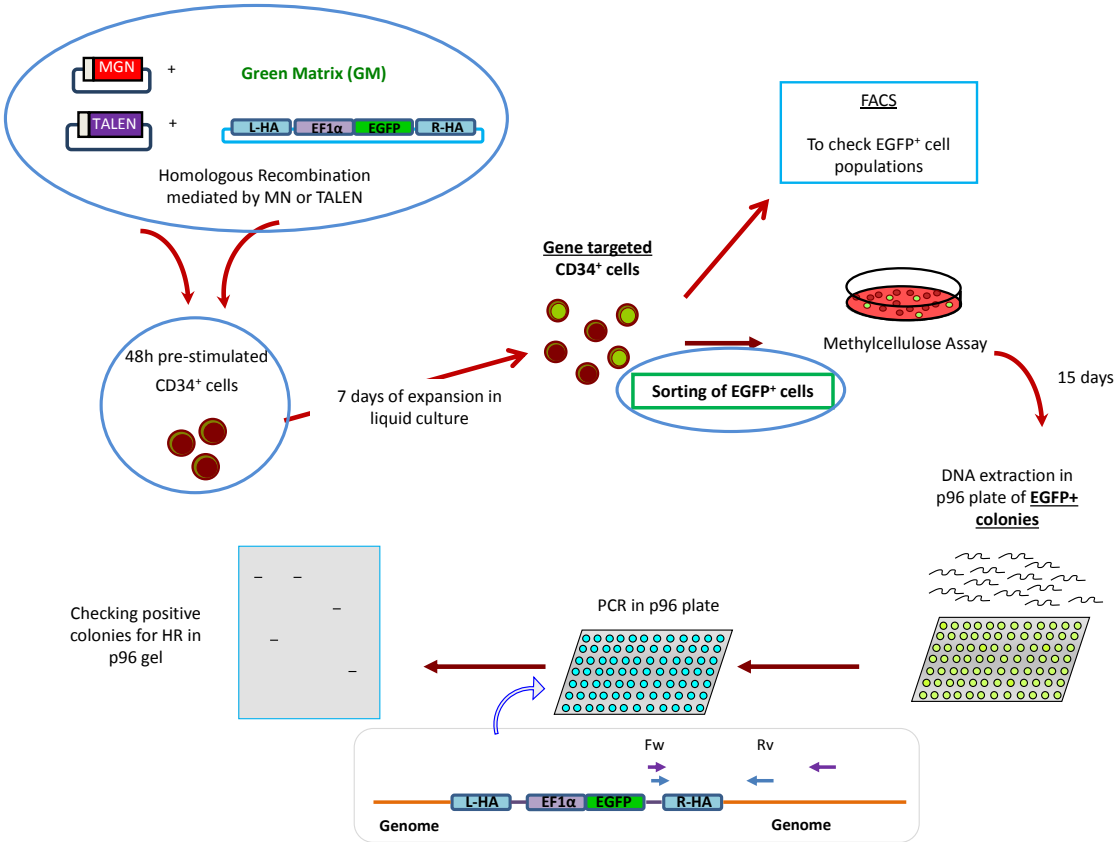

Figure 36: Scheme of the protocol used in CD34$^+$ Hematopoietic Progenitor Cells to induce HR mediated by SH6-nucleases.

- **Evaluation of the maximum amount of DNA to be nucleofected in CD34$^+$ hematopoietic progenitor cells**

In order to evaluate the toxicity and the transfection efficiency of different amounts of DNA in CD34+ cells, 2, 5, 10 and 15 µg of an EGFP/Cherry plasmid or 2 µg of a control pmaxGFP plasmid were nucleofected into the cells. The viability and the percentage of EGFP$^+$ cells were measured by FACS 48 hours after nucleofection.

Cells nucleofected with 2 µg of the control of DNA (pmaxGFP) showed a viability around 70%, (**Figure 37A**) but in the rest of the conditions used the viability ranged between 33.6-51.1%, indicating that nucleofection with higher amounts of DNA reduced cell viability. The analysis of the fluorescence cells in the FACS showed a progressive increase up to 44.1% of fluorescent cells when the GFP/Cherry plasmid was used. **Figure 37B**.

These results indicated that even a large plasmid of 6,222 bp can be introduce into CD34$^+$ cells by nucleofection compatible with viabilities of 30-50%.

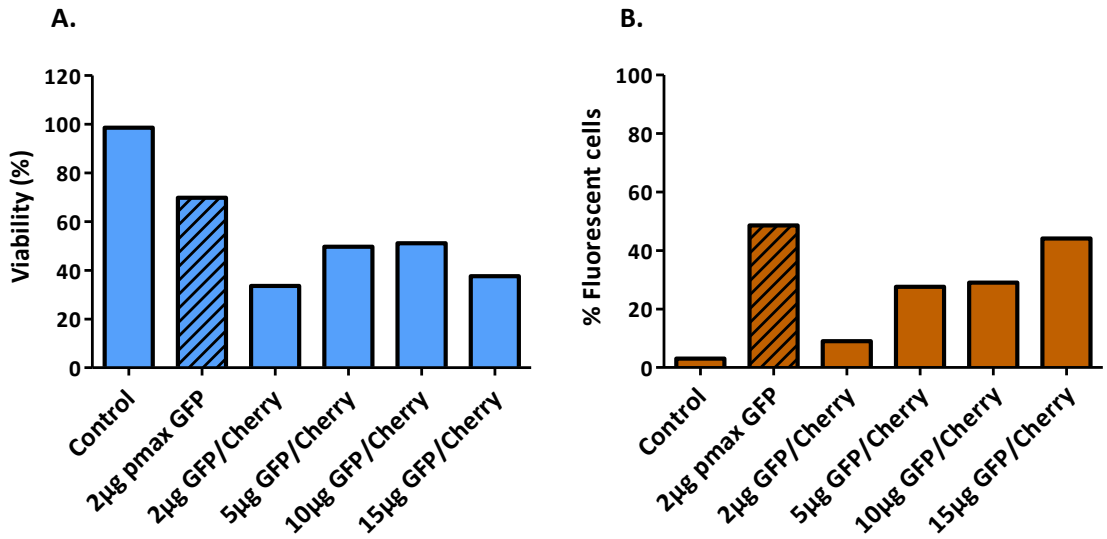

Figure 37: **A.** Viability of CD34+ cells (DAPI-) nucleofected with a control DNA (pmaxGFP, 3 kb) or with a large plasmid (6kb) carrying the EGFP and the Cherry reporter genes. **B.** Percentage of Fluorescent CD34$^+$ cells corresponding to samples analyzed in **A.**

- **CD34$^+$ hematopoietic progenitor cells maintain their clonogenic capacity after nucleofection and sorting of EGFP$^+$ cells with a control GFP/Cherry plasmid**

In subsequences experiments, CD34$^+$ cells were nucleofected with 2 µg of pmaxGFP or with 5.0 or 10.0 µg of the GFP/Cherry plasmid. Cells were grown in culture-media (see Materials and Methods) for seven days and then analyzed by FACS. As shown in **Figure 38A**, similar percentages of CD34+ cells, around 90%, were observed in all groups indicating that the

nucleofection process with or without DNA was not affecting the proportion of $CD34^+$ cells. Regarding the percentage of $CD34^+$ cells that expressed GFP, we observed that the highest percentage (49.2%) of $EGFP^+$ cells was obtained in cells nucleofected with 2 µg of pmaxGFP. GFP/Cherry nucleofected cells showed increasing percentages of GFP^+ cells as the GFP/Cherry was increase up to 25% of EGFP $^+$ cells with the dose nucleofected (**Figure 38A)**.

$CD34^+/GFP^+$ cells were then sorted and methylcellulose assays were performed to test the toxicity and the proliferation integrity after the nucleofection and the sorting analysis. As shown in **Figure 38B**, the nucleofection process, particularly when high doses of the GFP/Cherry plasmid were used, affected the number of colonies. No effects were observed however in samples nucleofected with the 2 µg of pmaxGFP. In all instances, myeloid and erythroid colonies were obtained.

These results indicated that although the nucleofection and sorting procedures reduced the proliferation integrity of $CD34^+$ cells, 25-40% of their clonogenic potential was preserved.

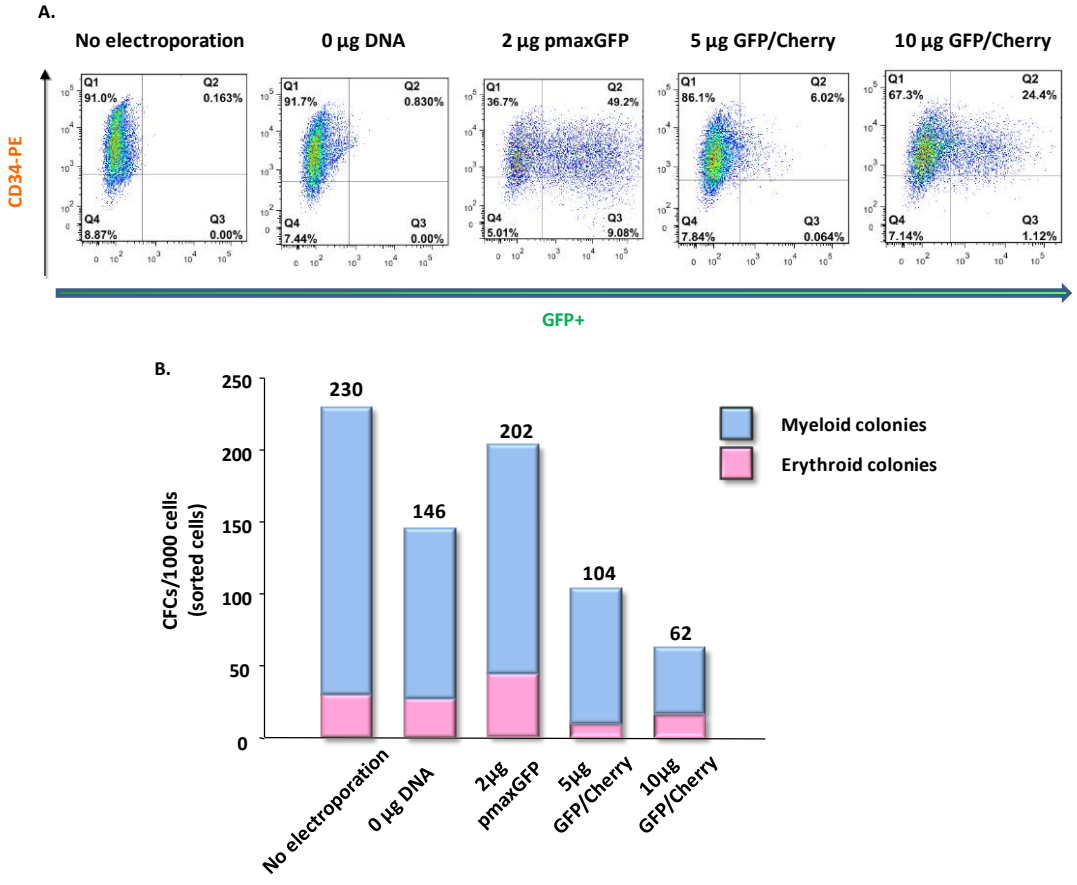

Figure 38: **A.** Percentages of CD34+ and GFP+ cells 7 days after nucleofection with the pmaxGFP or the GFP/Cherry plasmids **B.** Number of myeloid and erythroid colonies generated per 1,000 $CD34^+/GFP^+$ cells seeded in methylcellulose.

- **Comparison of the efficacy of the SH6v5-MN delivered as DNA or mRNA in CD34$^+$ cells**

To investigate the cytotoxicity induced by the SH6v5-MN as DNA or mRNA, CD34$^+$ cells were nucleofected with two doses of SH6v5-MN: 5 or 10 µg of DNA, or 10 or 20 µg of mRNA. Two days after nucleofection, the number of cells of each condition was determined and represented as the ratio with respect to the number of non nucleofected cells (**Figure 39A**). The toxicity of cells only suffering the pulse, or nucleofected with either 2 µg of the pmaxGFP or with 10 µg of mRNA GFP (a nucleofection control of the mRNA) was around 0.4 (**Figure 39A**). A higher toxicity was induced by SH6v5-MN used as DNA. Nevertheless, the nucleofection of CD34$^+$ cells with the mRNA induced a lower toxicity, as compared with the DNA MN.

Even more, the mRNA nucleofection seems to be more effective to transfect a most uniform cell population, as shown in **Figure 39B**, in which GFP mRNA control shows a concentrate population compared to the pmaxGFP, in which the intensity of the EGFP$^+$ cells is more variable.

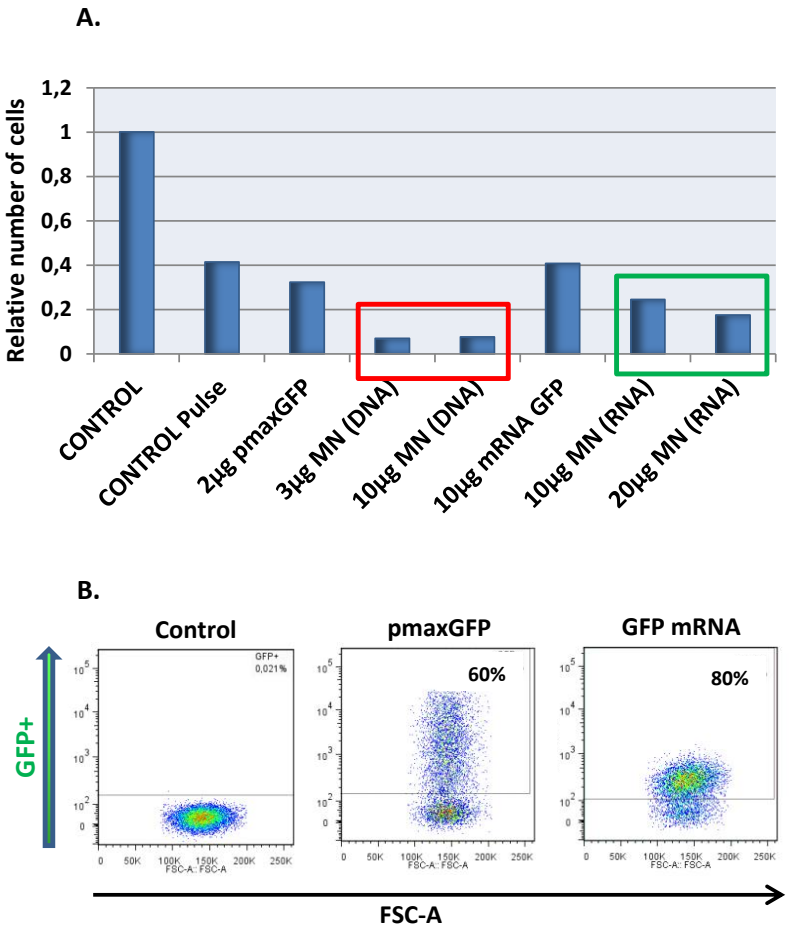

Figure 39: **A.** Survival rate of CD34+ cells after nucleofection with the nucleases delivered as DNA or mRNA. Columns inside the red square correspond to conditions nucleofected with DNA, and inside the green square, the conditions with the mRNA, which shows higher viability. **B.** The panel shows the controls of nucleofection, the pmaxGFP and mRNA of an EGFP reporter gene.

Once we determined that the SH6v5-MN was less toxic when used as mRNA vs DNA, we tested if the mRNA SH6v5-MN was functional. The Surveyor assay shown in **Figure 40** that the delivery of nucleases as DNA or as mRNA, induced cleavage in the SH6 site, and more markedly using the mRNA SH6v5-MN, represented by the higher intensity of the bands of 179 and 171 bp **Figure 40**.

These results indicate that the as mRNA SH6v5-MN is bot less toxic and more efficient when compared with the same MN used as DNA plasmid.

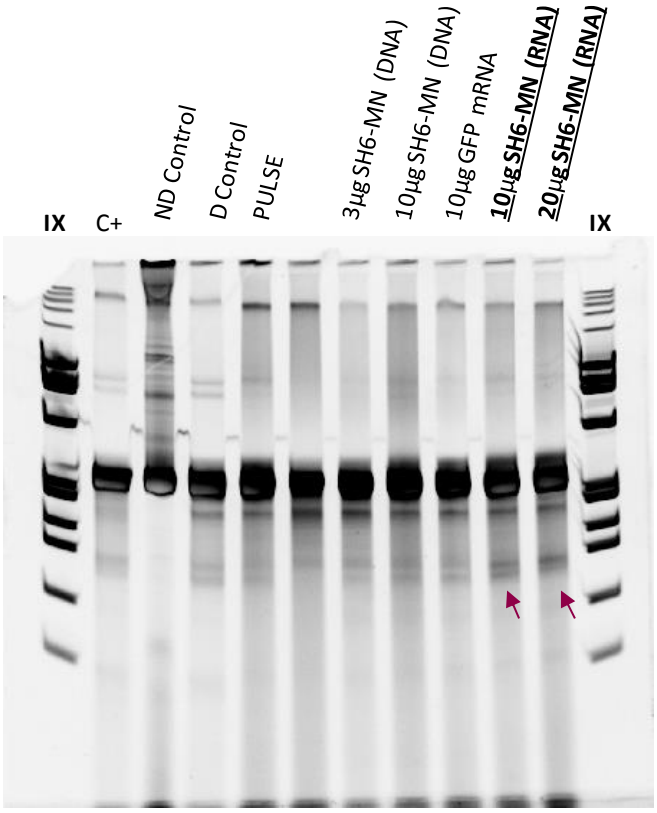

Figure 40: Surveyor assay analyses corresponding to the SH6v5-MN delivered as DNA or mRNA in CD34+ cells. The red arrows mark the more intense bands obtained in conditions nucleofected with the MN as mRNA.

- **Selection of the pre-stimulation period required to activate the cell cycle status of the CD34$^+$ cells**

Homologous Recombination is reported to be more efficient when cells are in the S/G2 phases of the cell cycle. In order to improve the efficiency of HR in CD34$^+$ cells, cells were thawed and cultured in pre-stimulation media, and cell cycle was checked by FACS at 24, 48 and 72 hours of culture. As represented in **Figure 41**, after 48 hours in culture there was an increase of the percentage of cells in G2 and S phases compared to that observed at 24 and 72 hours. In view of these results, experiments of HR were performed in CD34$^+$ cells after 48 hour of pre-stimulation.

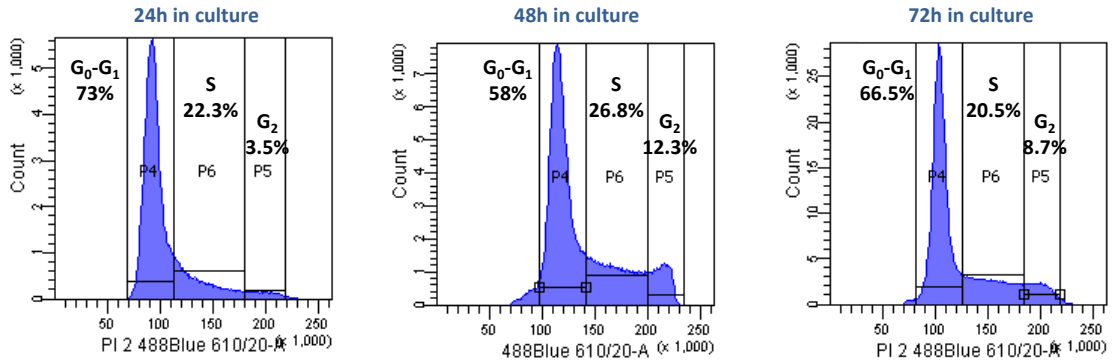

Figure 41: Pictures showing the proportion of CD34$^+$ cells in G2 and S phases of the cycle after 24, 48 or 72 hours of pre-stimulation.

- **Analysis of the HR events in CD34+ hematopoietic progenitor cells subjected to EGFP$^+$ cell sorting**

Homologous Recombination experiments were performed using the SH6v5-MN following the scheme represented in **Figure 36**, accordingly to the steps described in the following points.

Thawed CD34+ cells were cultured in pre-stimulation media for 48 hours. One million of cells per condition were then nucleofected with 4 µg of the GM alone or with the SH6v5-MN as DNA or as mRNA, or with the SH6-TALEN. Cells were maintained in culture-media, and the following analyses were performed.

- **Kinetics of CD34+/EGFP+ cells along the incubation of nucleofected cells**

In all instances most of the incubated cells remained positive for CD34$^+$ expression in nucleofected cells. When EGFP analyses were conducted, similar results were obtained in all samples nucleofected with the GM, regardless of the presence of the nucleases. These results thus indicate that the EGFP fluorescence observed up to 4 days after nucleofection should be mainly due to the transient expression of non-integrated copies of the GM.

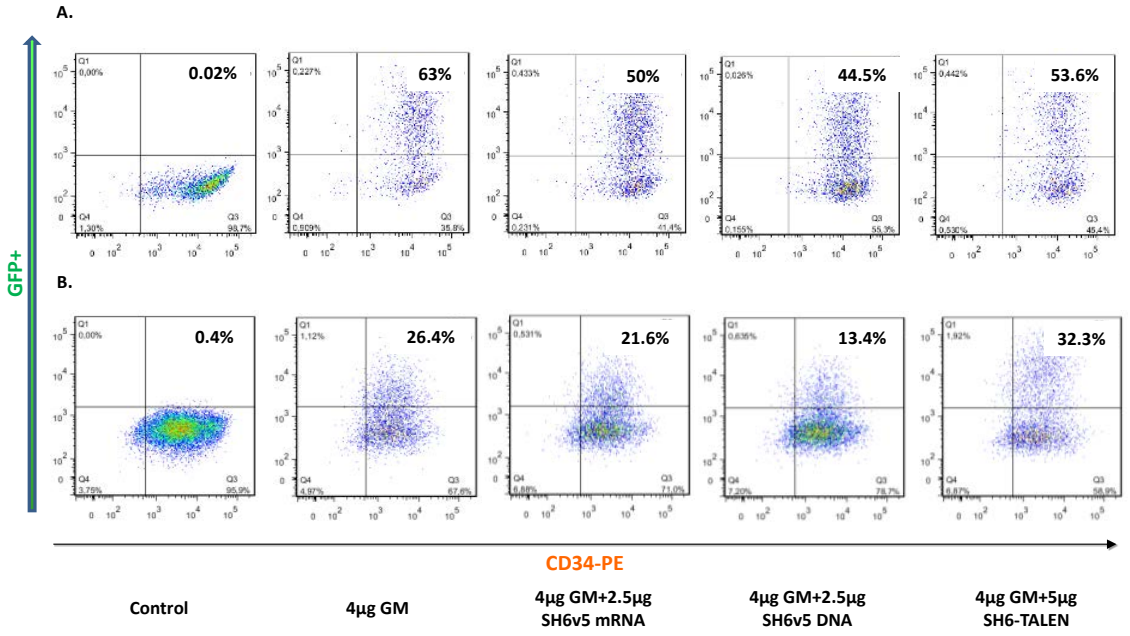

Figure 42: Flow cytometry analysis of CD34+ and EGFP+ cells after 48 hours (**A.**) or 4 days post-nucleofection **B.** with the GM and the different nucleases.

- Analysis of the clonogenic potential of EGFP$^+$ nucleofected cells

We also evaluated the clonogenic capacity of nucleofected cells. For this purpose, at 7 days post-nucleofection cells were counted and seeded in methylcellulose. Additionally, EGFP+ cells (4%-9%) were sorted from each condition and seeded in methylcellulose cultures as previously performed (**Figure 43**).

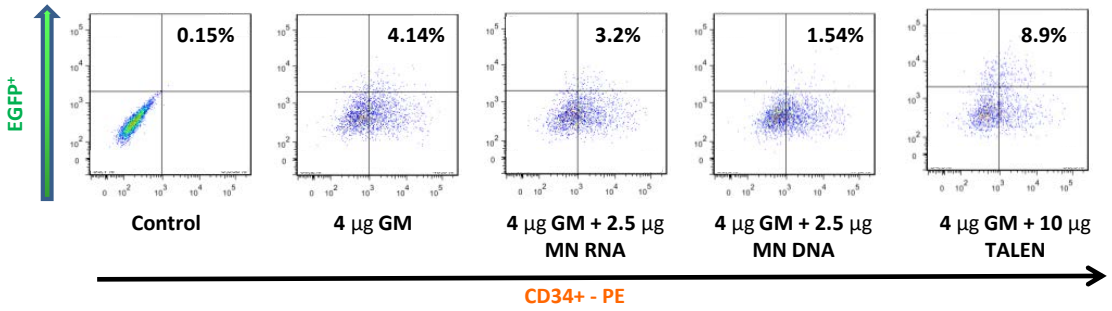

Figure 43: Analyses of the percentage of CD34+/EGFP+ cells, 7 days after nucleofection.

After 14 days in culture, the number of colonies generated in each condition was scored (**Figure 44A**). When the proportion of colonies per 10^5 cells was considered, we observed that the cell sorting reduced the number of colony forming units. The nucleofection of the nucleases seemed not to affect this number. When the number of EGFP+ colonies was scored

in post-sorted cells, we realized that no differences were observed but for the mRNA MN condition in which this number is lower (**Figure 44B)**.

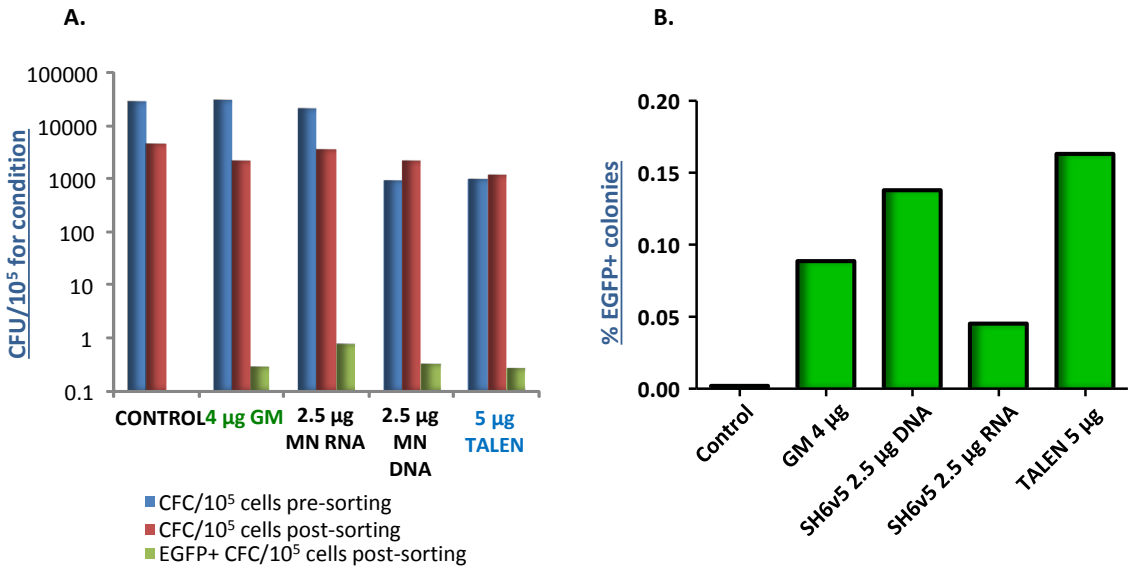

Figure 44: **A.** Number of CFU/10^5 cells obtained after 14 days in methylcellulose cultures. **B**. Percentage of EGFP+ colonies among the total number of colonies obtained.

- **Analyses for specific integration in SH6 locus**

Nested-PCR was used to demonstrate that EGFP+ colonies carried a specific integration of the GM in the SH6 locus (see materials and methods). EGFP+ colonies were picked and their DNA extracted and Nested-PCR analyses were performed. Pseudoclones obtained from our previous experiments in HEK-293H cells were used as positive controls. The presence of a band of 1,571 bp, would be indicative of specific integrations of the GM in the SH6 site.

Strikingly, none of the 24 $EGFP^+$ colonies obtained in these experiments showed the presence of this band, **Figure 45**, indicating the absence of the specific integration of the GM in the *SH6* locus.

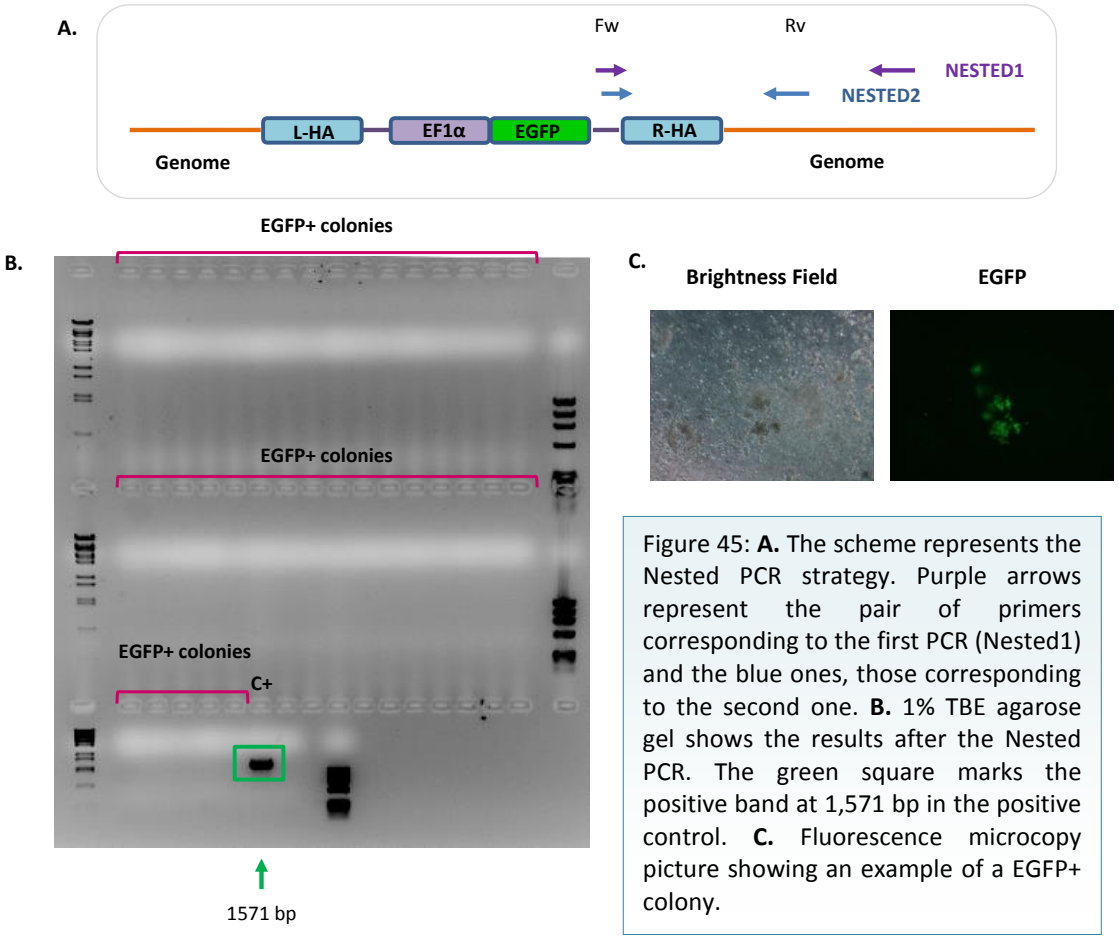

Figure 45: **A.** The scheme represents the Nested PCR strategy. Purple arrows represent the pair of primers corresponding to the first PCR (Nested1) and the blue ones, those corresponding to the second one. **B.** 1% TBE agarose gel shows the results after the Nested PCR. The green square marks the positive band at 1,571 bp in the positive control. **C.** Fluorescence microcopy picture showing an example of a EGFP+ colony.

4.2 Homologous recombination experiments in the absence of EGFP$^+$ cell sorting using SH6-TALEN and SH6v5-MN in CD34+ hematopoietic progenitors

Data obtained in experiments shown in **Figure 45**, together with either observations from our collaborators showing that a high expression of the EGFP could be a consequence of the integration of more than one copy of the reporter gene in the SH6 site, or due to random integrations (Eyquem, Poirot et al. 2013), made us to consider the convenience of performing a new experimental protocol in which nucleofected cells expressing EGFP were not sorted.

After 48 hours of pre-stimulation, one million of CD34+ cells were nucleofected per condition and maintained in culture-media. The proportion of EGFP+ cells was measured 2 days post-nucleofection by FACS and primary populations was evaluated. Cells were counted and aliquots were seeded in methylcellulose plates. Additional aliquots were used to extract the DNA to also analyze by PCR for the presence of HR events. Colonies generated after 14 days in methylcellulose were counted, picked and the DNA extracted to perform Nested-PCR analyses to search for specific integration of the GM (**Figure 46**).

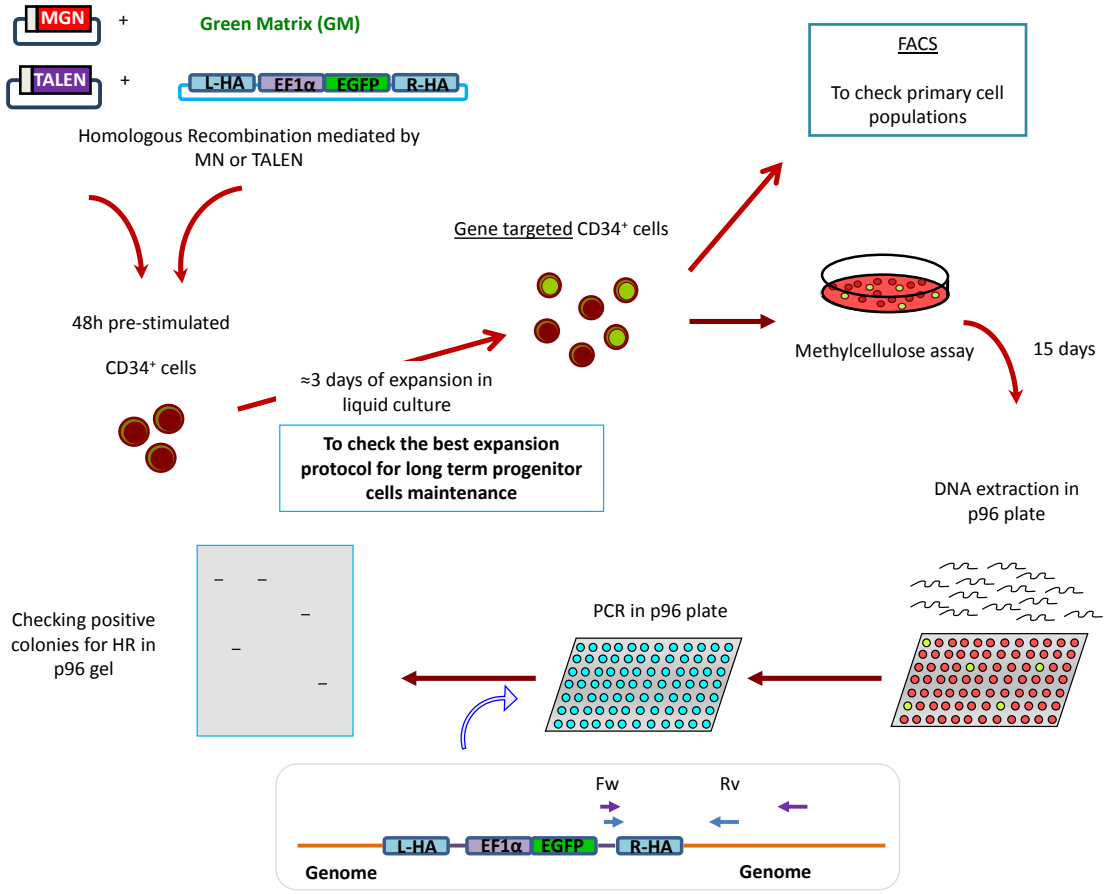

Figure 46: Scheme of the new protocol to promote nucleases mediated HR in the SH6 site in CD34$^+$ hematopoietic progenitor cells.

- **Analysis of HR in CD34$^+$ hematopoietic progenitor cells not subjected to pre-EGFP$^+$ cell sorting**

One million of pre-stimulated CD34$^+$ cells were nucleofected with 4 µg of the GM alone or either with 2.5 µg of SH6v5-MN, used as DNA or mRNA, or with 2.5 or 5 µg (total amount of both subunits) of SH6-TALEN as DNA. Cells were maintained in the culture media and analyzed as described in **Figure 46**.

- **Viability and transfection efficiency**

In order to evaluate the toxicity of the nucleases the viability of nucleofected cells was measured by FACS at 4 days post-nucleofection. As shown in **Figure 47**, all the nucleofected conditions showed a slightly lower viability compared to Control non-nucleofected cells.

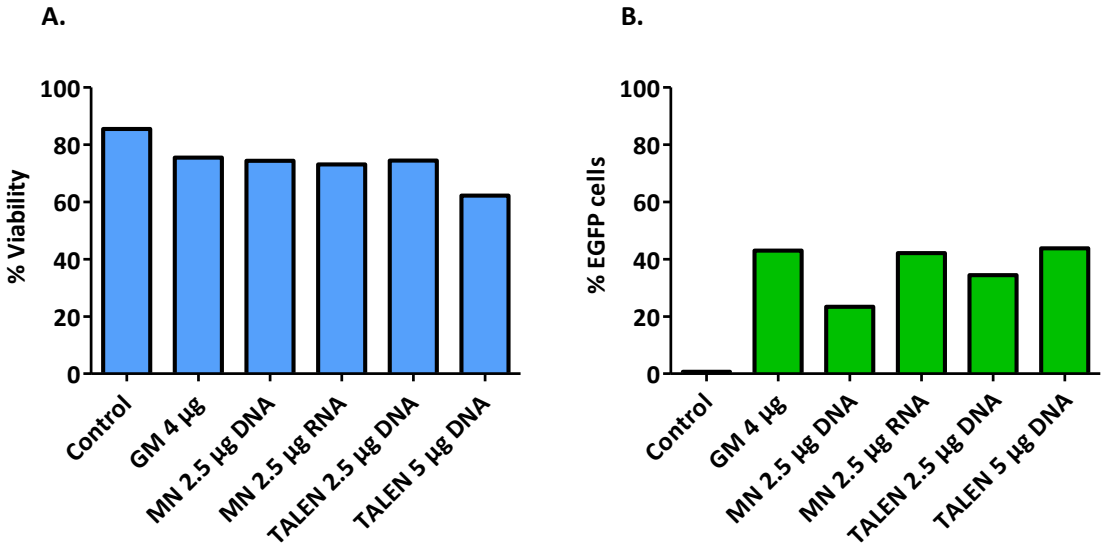

Figure 47: **A.** Viability of the cell culture (DAPI negative cells) at 4 days post-nucleofection measured by FACS. **B.** Percentages of EGFP$^+$ cells determined at 4 days post-nucleofection.

With respect to the EGFP expression, nucleofected cells showed 20- 40% of EGFP expressing cells. These percentages were similar to the ones obtained in previous experiments, indicating the consistency of our experimental protocol to electroporate CD34$^+$ cells

These results show that the electroporation of CD34+ cells with the nucleases and the GM was slightly toxic for the nucleofected cells, with an acceptable delivery and expression of the transgene.

- **Maintenance of the stem cell primitiveness**

To analyze if our cells were affected by the different manipulations (thawing, nucleofection, DNA delivery, expansion…), we measured by FACS different stem cell subpopulations following the protocol proposed by Dr. J. Dick (Doulatov, Notta et al. 2012). We also analyzed whether the GM was more efficiently transfected in primitive or in differentiated cells.

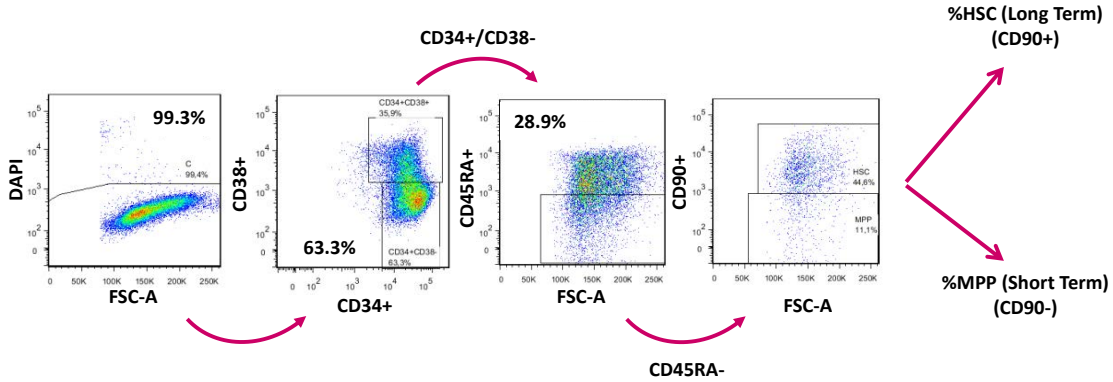

Figure 48: Scheme of the FACS analyses to evaluate the HSPC compartments according to the protocol proposed by Doulatov *et al.* (Doulatov, Notta et al. 2012).

As shown in **Figure 48**, we identified two different HSC subpopulations in our electroporated samples. The most primitive with a long term repopulating ability (Hematopoietic Stem Progenitor Cells) [HSPC] ($CD34^+/38^-/45RA^+/CD90^+$) and Multipotent Progenitors (MPP) ($CD34^+/38^-/45RA^+/CD90^-$), with short-term repopulating ability.

With respect to the long term repopulating cells, percentages of 5 to 10%, were observed, similar than those observed in the control non-electroporated sample (14%). EGFP+ HSPC percentages were among 2-4% of the total population and, depending on the condition, 16% to 46% of the HSPC cells were EGFP+, indicating that a significant proportion of the HSCs carried the GM. The percentage of MPPs present in the nucleofected cells was lower, as compared to the HSPCs (3-6%). However, the proportion of EGFP+ cells within this population was similar to the proportion of EGFP+ cells found in the HSPC fraction. These results indicated that our experimental conditions were efficient to preserve the stem capability of the culture.

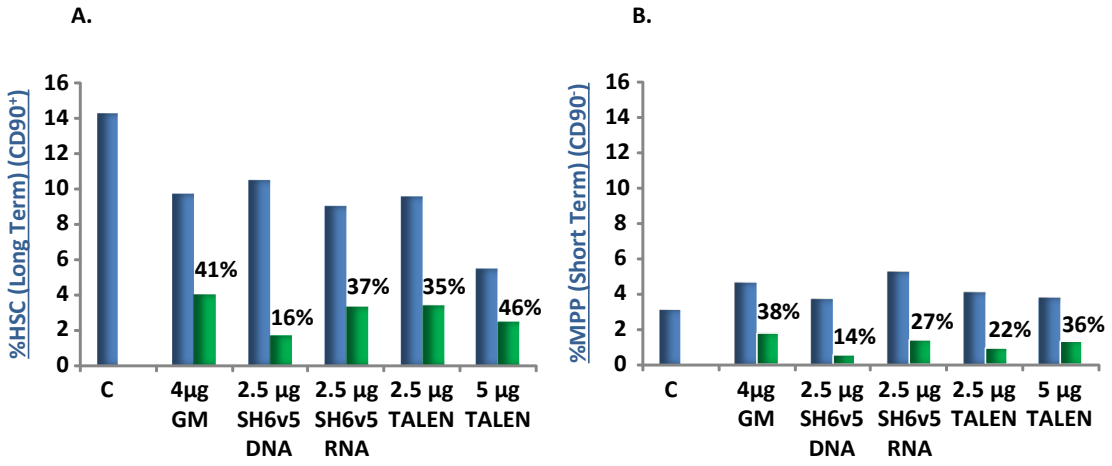

Figure 49: **A.** Percentages of HSC (blue bars) and EGFP+ HSC (green bars) in the pool of cells, 4 days post-nucleofection. Percentages above the green bars mean the percentage of HSC that were EGFP+. **B.** Percentages of MPP (blue bars) and EGFP+ MPP (green bars), and the percentages of MPP that were EGFP+.

- **Study of the specific integration of the GM in *SH6* locus in the pool of nucleofected CD34+ cells and in methylcellulose colonies**

Four days post-nucleofection, cells were counted and part of the cells was seeded in methylcellulose assays to generate hematopoietic colonies. Picked colonies and the pool of nucleofected cells were analyzed by Nested PCR.

As shown in **Figure 50**, in the total nucleofected population a band of 1,571 bp, corresponding to the specific integration of the GM in the SH6 site was observed in most cases, even in the absence of the nuclease.

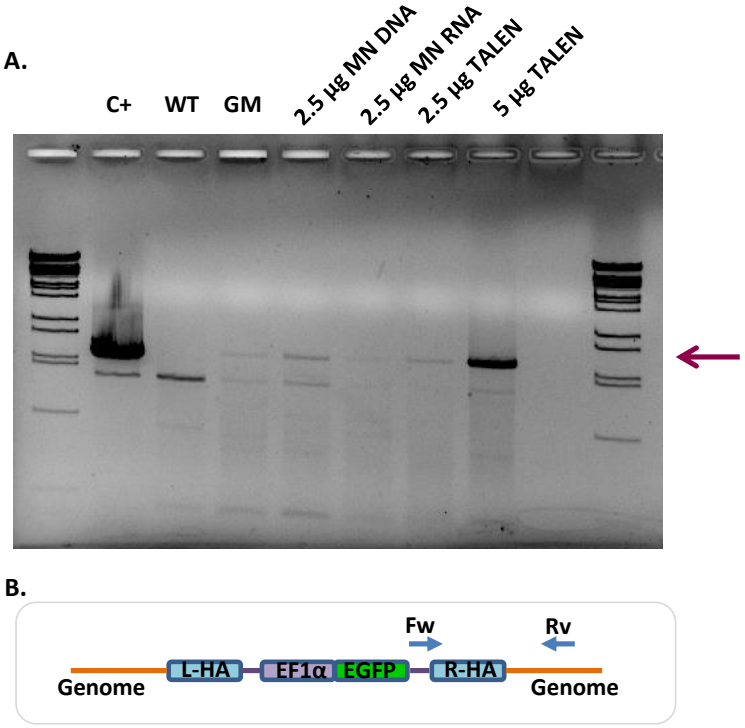

Figure 50: **A.** TBE agarose gel shows the 1,571 bp band, marked with the red arrow, indicative of the specific integration of the GM in SH6 locus. The most intense band corresponds to the 5 µg of SH6-TALEN. **B.** the scheme shows the binding sites of the pair of primers used for the specific PCR for the SH6 site.

After 14 days, colonies from methylcellulose assays were counted and around 200 colonies per condition were picked, DNA extracted and analyzed by Nested-PCRs. The number of total colonies generated in each condition nucleofected was similar with the exception of the condition in which 5 µg of TALEN was used, in which a lower number of colonies were obtained (**Figure 51A)**.

As shown in **Figure 51B**, when we performed the Nested-PCR in picked colonies, only one colony that corresponded to the 5 µg of SH6-TALEN condition was positive. This indicated the presence of a specific integration in SH6 locus in one primary hematopoietic progenitor cell, which corresponds to a percentage of HR of 2.1%.

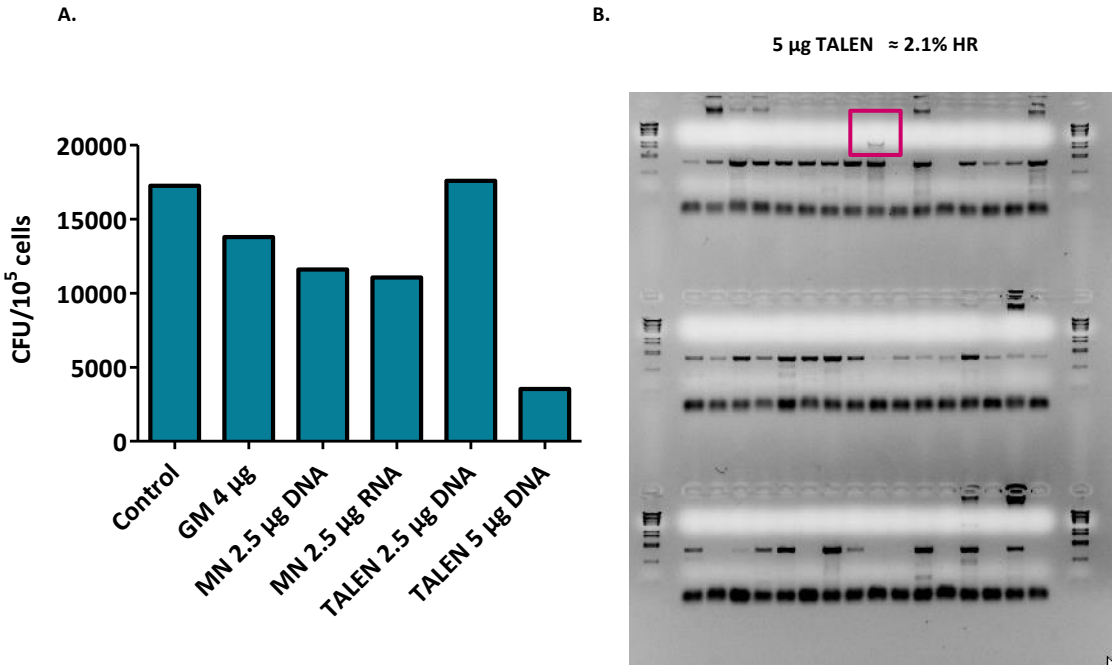

Figure 51: **A.** Number of colonies/10^5 cells generated after nucleofection with the conditions shown in **Figure 50**. **B.** The agarose gel electrophoresis displays the result of the Nested-PCR which shows a positive band inside the purple square.

- **mRNA synthesis of the SH6-TALEN by cloning an 3´UTR from the β-globin gene**

In order to limit the toxicity induced by the nucleofection of the 5 µg of the SH6-TALEN, the only condition in which targeted integration took place in CD34+ cells, we aimed to synthesize this as mRNA nuclease. This synthesis could be difficult because of the long size of this nuclease. For this reason we cloned the 3´UTR from the β-globin gene in our plasmid, as it has been reported that this sequence stabilizes the RNA of the β-globin gene.

In order to test if the SH6-TALEN mRNA was active, CD34+ cells were nucleofected with the TALEN either as DNA or as mRNA either harboring or not the 3´UTR (**Figure 52A)**. As shown in **Figure 52A**, the viability of the cells was increased up to 3 times (around 70% of viability) compared to the DNA-TALEN condition. Moreover, the surveyor assays showed that the percentage of insertions-deletions (INDELS) generated by the mRNA TALEN was always higher compared to the DNA TALEN.

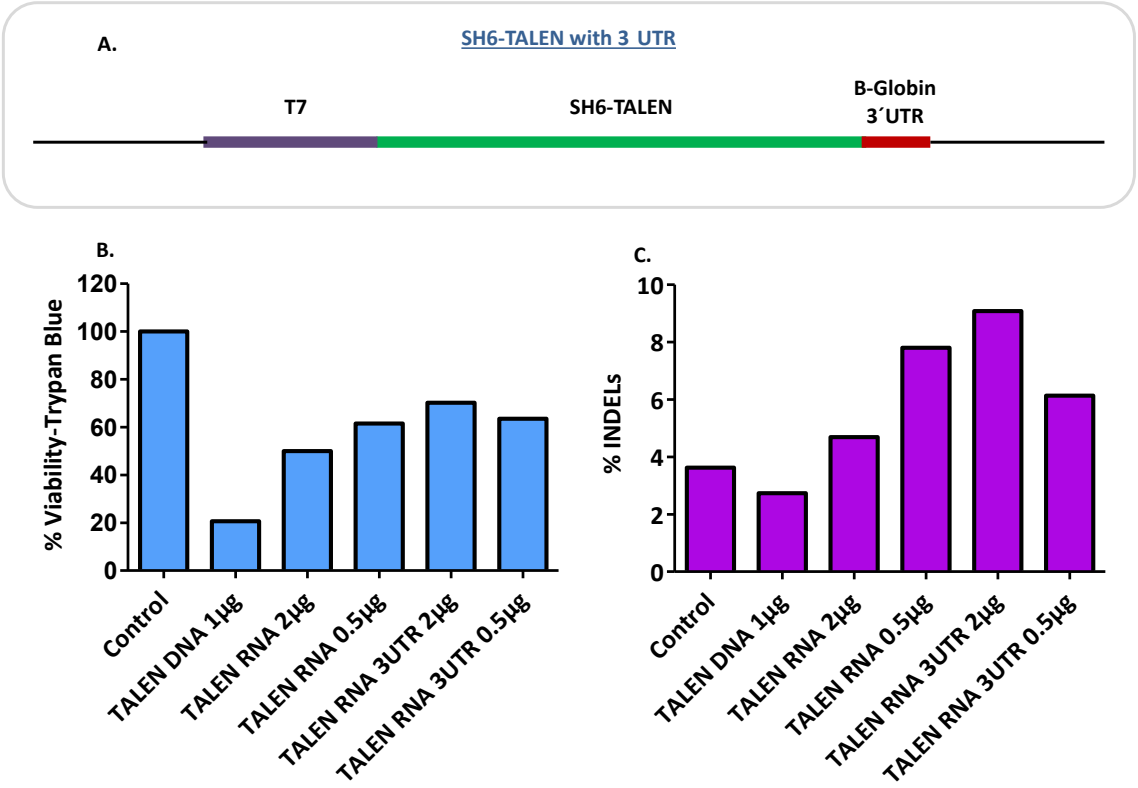

Figure 52: **A. S**cheme of the plasmid used to synthesize the mRNA from the TALEN. The T7 promoter drives the synthesis of the SH6-TALEN. The 3´UTR structure is expected to stabilize this mRNA. **B.** Viability of nucleofected CD34+ cells with the TALEN as mRNA and as DNA. **C.** Efficiencies of the nucleases delivered in different configurations.

These results allowed us to conclude that the use of the SH6-TALEN synthetized as mRNA with the 3´UTR was less toxic and more active in CD34+ hematopoietic progenitors. Thus we decided to use this 3´UTR mRNA TALEN in the next experiments.

• Analysis of the efficacy of the nucleofector device to mediate HR in CD34^{+} cells with MNs and TALEN

While in all the previous experiments we used the AMAXA Device I, in a new set of experiments we investigated the toxicity and nucleofection efficacy of the new AMAXA nucleofector device 4D instead of the device I used in previous experiments.

- Comparison of nucleofection efficiency between Strips and Cuvettes

In order to choose the more efficient approach, we performed a comparative analysis between the strips and the cuvettes used in the 4D device. One million of CD34+ cells were

nucleofected in the cuvettes, with 2 µg of the nucleofection Control pmaxGFP, and increasing doses of the donor GM (4, 5 and 8 µg). 100,000 of CD34+ cells were nucleofected in the strips with 1 µg of pmaxGFP or 0.8, 1.0 or 1.6 µg of the GM.

When we compare the viability obtained in cells nucleofected in the cuvettes and strips, higher survival rates were obtained with the strips (**Figure 53A)**. With respect to the percentage of EGFP+ cells, similar percentages (60-66%) were obtained when the pmaxGFP was used. When the GM was used, the percentages were higher for the cuvettes (42-50%) in contrast to the range obtained with the strip (7-12%) (**Figure 53B**).

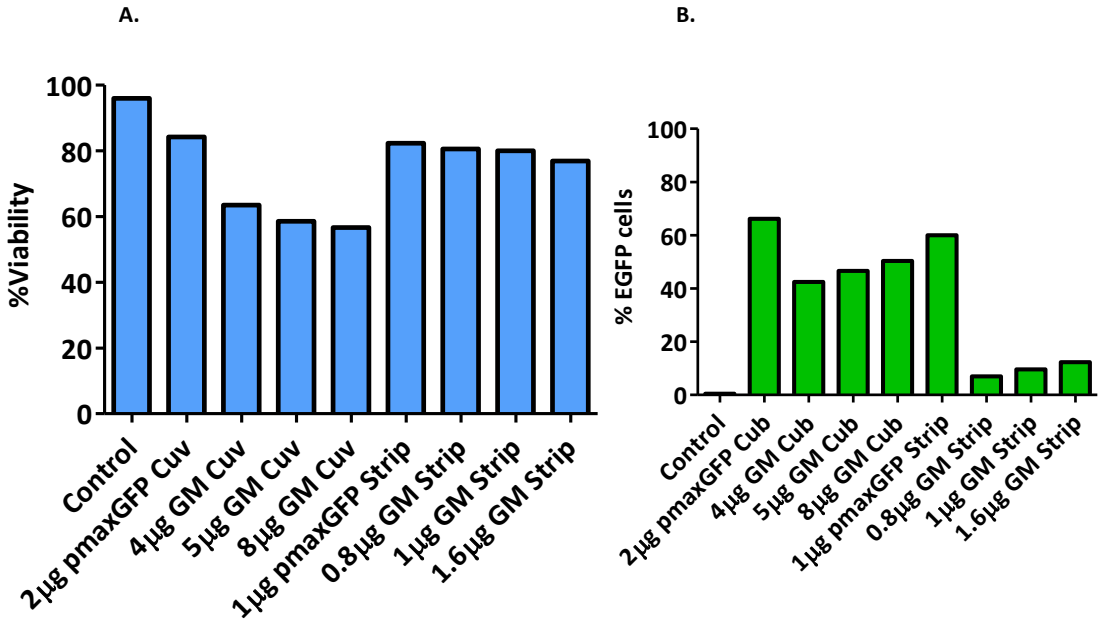

Figure 53: **A.** Cell viability (DAPI-) comparison between cuvettes and strips platforms. Cytometry analyses were performed 2 days post-nucleofection. **B.** Nucleofection efficiency comparison between, the two platforms.

With respect this result, we decided to perform the following experiments using the cuvettes.

.

- **Homologous Recombination experiments using Cuvettes: increased nucleofection efficiency**

Based on data showing higher percentages of $EGFP^+$ cells when the cuvettes were used, we performed additional experiments using SH6-TALEN and the GM.

One million of CD34+ cells per condition were nucleofected with GM alone or with 5 µg of the SH6-TALEN as total plasmidic DNA or total mRNA.

Two days after nucleofection, the viability of the cells was around 40% in all nucleofected conditions (**Figure 54A).** Also similar percentages of EGFP+ cells (49 to 54%) were observed (**Figure 54B)**.

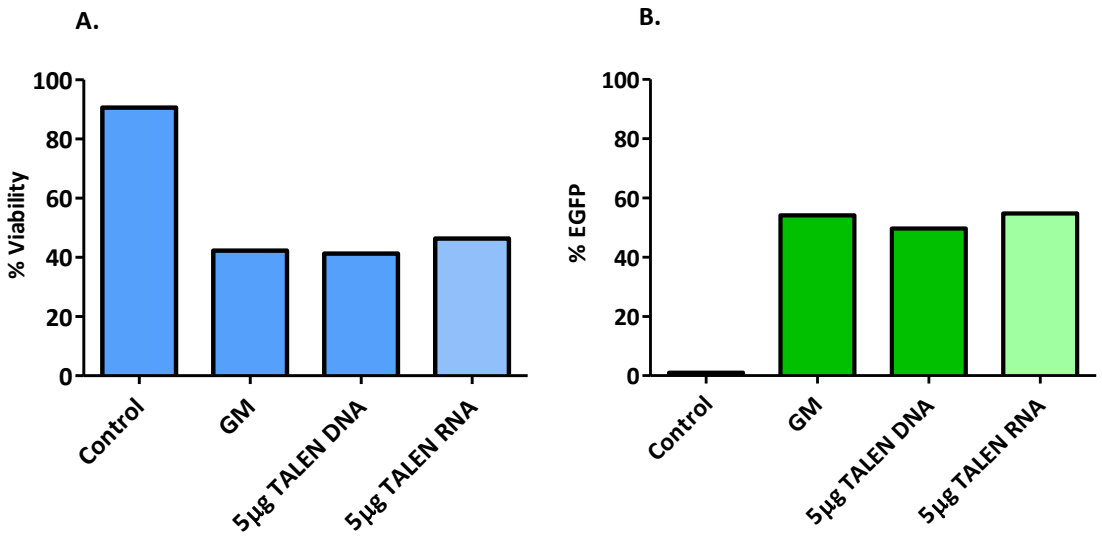

Figure 54: **A.** Viability of the cells 2 days post-nucleofection. **B.** EGFP+ percentages are shown.

The phenotype of the HSPCs at days 2 and 7 after nucleofection was analyzed by FACS. Similar results to the previous experiments (**Figure 49**) were obtained, showing a higher HSC population compared to MPP. Both of them contained subpopulation that expressed the EGFP transgene, (Data not shown). These results indicate that most of the primary cell carried the GM and suggested that a high percentage of them could have it integrated in the genome.

Nucleofected cells were also plated in methylcellulose assays and after 14 days colonies were picked and the DNA was extracted and analyzed for specific integrations in the *SH6* locus by Nested-PCR. Focusing in the number of colonies, similar results were obtained compared with previous experiments (**¡Error! No se encuentra el origen de la referencia.**), since no differences in the number of colonies among all the nucleofected conditions were generated. The number of colonies in nucleofected cells was always lower compared with the control, revealing a negative effect of the nucleofection with either of DNA or mRNA in the progenitor cells.

Unexpectedly, none of the colonies showed the specific integration of the GM in *SH6* locus.

Although the nucleofection efficiency obtained with the cuvettes was similar to that of previous experiments performed in the device I, no gene targeting of the GM in the SH6 site was obtained.

In a new experiments, one million of $CD34^+$ cells were nucleofected with increasing doses of the SH6-TALEN as total DNA or total mRNA (with or without the 3´UTR) in order to force the delivery of the nucleases as an attempt to improve the HR efficiency.

One day post nucleofection, the viability and percentage of EGFP+ cells were analyzed by FACS. The viability of the cells was around 40-60% for all the nucleofected conditions. Similar

lower percentages of EGFP$^+$ cells (10-20%) were obtained in all the nucleofected conditions (**Figure 55A)**.

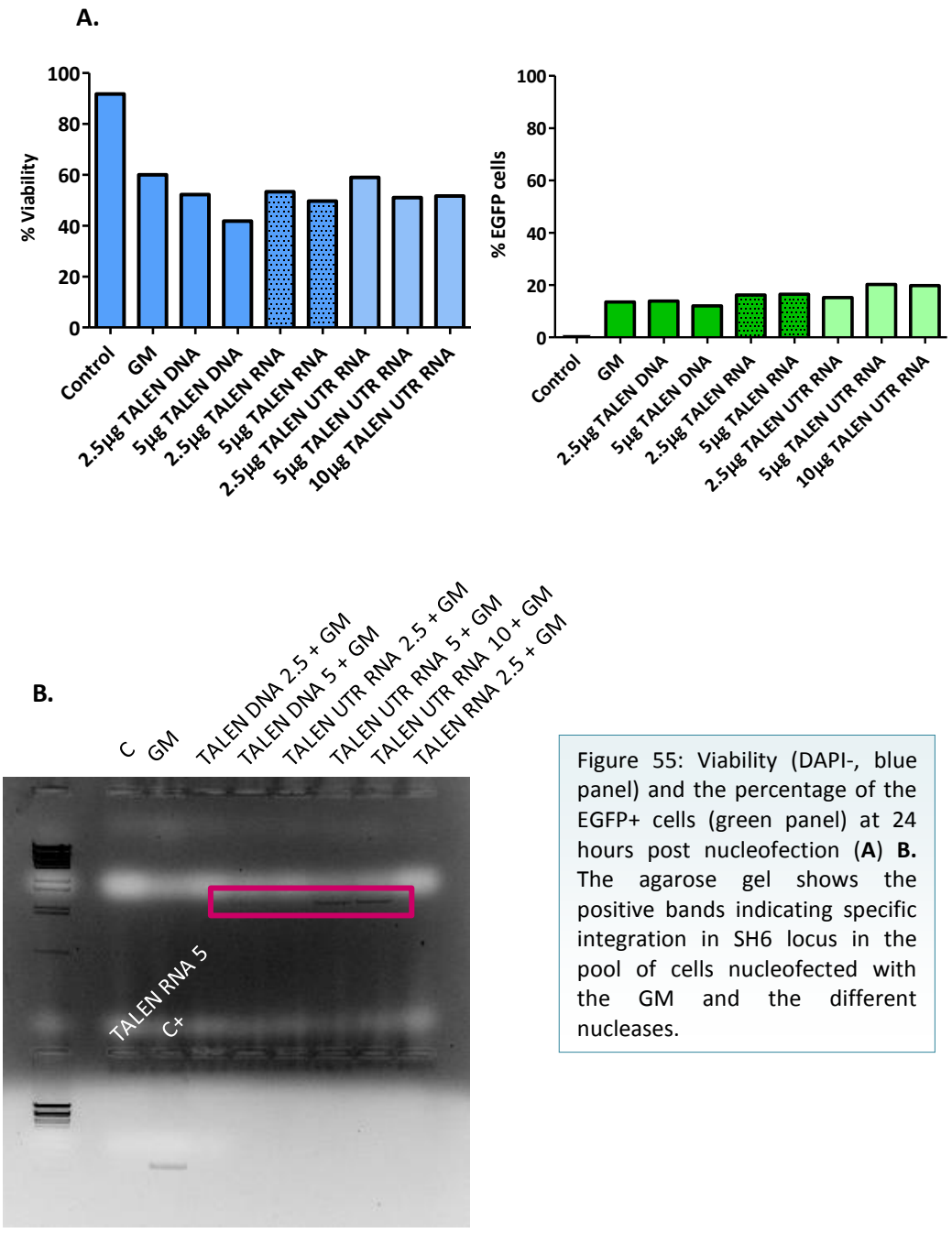

Figure 55: Viability (DAPI-, blue panel) and the percentage of the EGFP+ cells (green panel) at 24 hours post nucleofection (**A**) **B.** The agarose gel shows the positive bands indicating specific integration in SH6 locus in the pool of cells nucleofected with the GM and the different nucleases.

DNA from the pool of cells was analyzed by specific PCR and positive bands were obtained in the 5 µg DNA TALEN and the 3`UTR mRNA TALEN conditions. No band was obtained neither in GM alone nor in the no 3´UTR-mRNA condition **Figure 55B**.

The HSPCs phenotype present at day 2 after nucleofection showed a higher HSC population compared to the MPP. Both of the subpopulation showed a low percentage of EGFP+ cells, as already observed in the previous experiments (Data not shown) (**Figure 49**).

Nucleofected cells were also plated in methylcellulose assays and after 14 days colonies were scored, picked and the DNA was extracted and analyzed for specific integrations in SH6 locus by Nested-PCR, 200 CFUs per condition. Again, none of these colonies showed the specific integration in SH6 locus (Data not shown). Focusing in the number of colonies, similar numbers were observed among all the nucleofected conditions (Data not shown).

Taken together, the results obtained with the Amaxa 4D indicated that this device was less efficient than the AMAXA Device I mediating HR in $CD34^+$ cells, at least with the MNs and TALEN used for the integration of the GM in the SH6 locus (**Figure 51**).

- **Improvement of the Homologous Recombination of the GM in the SH6 locus of the $CD34^+$ cells using AMAXA I nucleofector system**

Because de 4D device, either with strips or cuvettes used did not improve the efficiency of HR as compared with the AMAXA I, new HR experiments with the AMAXA Device I were conducted again with the GM and the SH6-TALEN used as mRNA or as DNA.

One million of CD34+ cells were nucleofected in each condition with 4 µg of the GM either alone, or with 10 µg of the SH6-TALEN as 3´UTR mRNA or with increasing doses of the TALEN as DNA, 2.5, 5 or 10 µg.

At 2 days post nucleofection the viability and the percentage of $EGFP^+$ cells were measured by FACS. Data in **Figure 56A** showed a progressive decrease in the cell viability as the dose of DNA was increased, even when the mRNA TALEN was used. Percentages of EGFP+ cells were similar and around 60% (53.4% to 75.3%) for all the conditions (**Figure 56B)**. DNA from the pool of cells was analyzed by specific Nested-PCR and a positive band was obtained for all the conditions in which cells were nucleofected with nucleases, indicating the presence of HR events in the pool of cells. No band was obtained in non nucleofected cells (Control) or in cells only nucleofected with the GM.

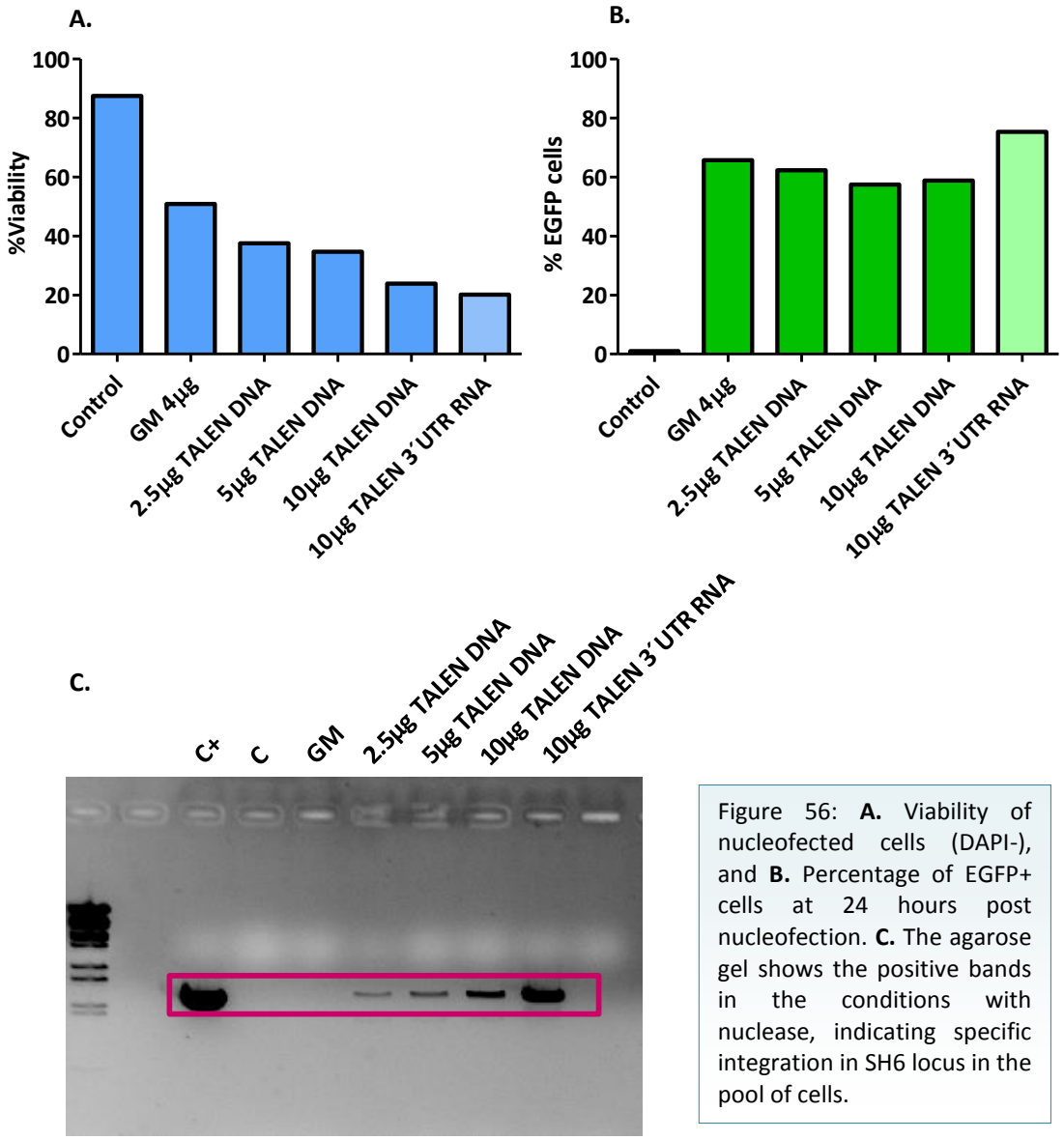

Figure 56: **A.** Viability of nucleofected cells (DAPI-), and **B.** Percentage of EGFP+ cells at 24 hours post nucleofection. **C.** The agarose gel shows the positive bands in the conditions with nuclease, indicating specific integration in SH6 locus in the pool of cells.

Nucleofected cells were plated in methylcellulose assays four days after nucleofection. After 14 days, colonies were scored, picked and the DNA was extracted and analyzed for specific integrations in the *SH6* locus by Nested-PCR.

Once again, a similar number of colonies was obtained in all the nucleofected conditions. (Data not shown)

With respect to the gene targeting, two positive colonies were obtained when cells were nucleofected with the 5 µg of the DNA TALEN, indicating a HR frequency of 1.04%. Six positive colonies among 192 were obtained when the TALEN dose was increased to 10 µg of DNA, indicating that the HR frequency was 3.13%. No HR was observed in the colonies when the TALEN was used a mRNA or at lower doses of DNA.

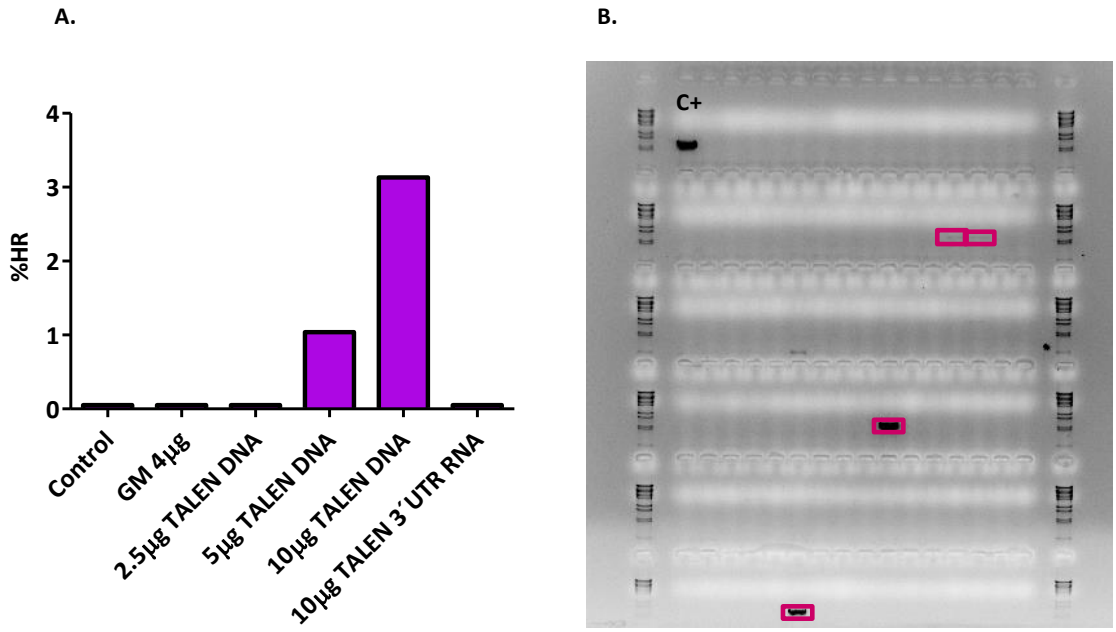

Figure 57: **A.** HR percentages induced by the SH6-TALEN using the GM alone or with the TALEN. **B.** Agarose electrophoresis gels showing positive bands, inside the purple boxes, for HR events from DNA obtained from CFUs.

Taken together, these final results indicate that AMAXA I device constitutes an adequate delivery method to induce HR in CD34+ cells, particularly when cells are nucleofected with the GM and with relative high doses of TALEN DNA (5-10 µg). The nucleofection did not affect the primitiveness of the HSPCs, or their clonogenic capacity.

Nuclease	Device	Platform	Viability (%)	EGFP$^+$ Population (%)	HR in cells cultured in liquid culture	HR in nucleofected colonies
MN DNA 0.1µg	4D	Strips	55	10	-	<1%
MN mRNA 0.1µg			60	10	-	<1%
TALEN DNA 1µg			50	10	-	<1%
TALEN RNA 3´UTR 2µg			55	10	-	<1%
TALEN DNA 2.5µg	4D	Cuvettes	50	15	No	<1%
TALEN DNA 5µg			55/40/40	15/55/10	No/ - /Yes	<1%/<1%/<1%
TALEN RNA 2.5µg			55	18	No	<1%
TALEN RNA 5µg			50	18	No	<1%
TALEN RNA 3´UTR 2.5µg			60	17	Yes	<1%
TALEN RNA 3´UTR 5µg			60/40/50	10/55/20	No/Yes/Yes	<1%/<1%/<1%
TALEN RNA 3´UTR 10µg			50	20	Yes	<1%

Table 9: Summary of the different nucleases and doses employed in HR experiments in CD34$^+$ cells using AMAXA Device 4D. The percentages of viability and EGFP+ population and the HR events induced are represented.

Nuclease	Device	Platform	Viability (%)	EGFP$^+$ Population (%)	HR Pool	HR CFUs
MN DNA 2.5µg	I	Cuvettes	75	20	Yes	<1%
MN RNA 2.5µg			75	40	Yes	<1%
TALEN DNA 2.5µg			75/40	40/60	Yes/Yes	<1%/<1%
TALEN DNA 5µg			60/35	40/58	Yes/Yes	2.1% 1%
TALEN DNA 10µg			20	60	Yes	3%
TALEN RNA 3´UTR 10µg			20	70	Yes	<1%

Table 10: Summary of the different nucleases and doses employed in HR experiments in CD34$^+$ cells using AMAXA Device I. The percentages of the viability and EGFP$^+$ cells and of the HR events induced are represented. The square includes the experiments positive for gene targeting in the SH6 locus with the GM in colonies.

DISCUSIÓN

Como se comentó en la introducción, la anemia de Fanconi es una enfermedad rara que se caracteriza por malformaciones congénitas, inestabilidad genética y predisposición al cáncer. Sin embargo, la principal causa de muerte en esta enfermedad es el fallo de médula ósea.

También se ha comentado que en los últimos años la terapia génica con el uso de LV ha tenido un gran desarrollo, habiendo sido aplicada exitosamente a diferentes inmunodeficiencias y otras enfermedades monogénicas. De hecho, un LV con el gen terapéutico *FANCA* ha sido desarrollado por nuestro laboratorio y aprobado como medicamento huérfano por la Comisión Europea, con el que se está llevando a cabo un ensayo clínico.

A pesar de todo esto, el principal problema de los LVs sigue siendo la posibilidad de generar fenómenos de mutagénesis y oncogénesis insercional. Es por esto que se siguen investigando nuevas estrategias para perfeccionar la inserción de los genes terapéuticos en sitios específicos del genoma, por ejemplo mediante el uso de nucleasas. Estas nucleasas, TALEN, meganucleasas, ZFNs y CRISPR/Cas9 son capaces de generar dobles roturas en el ADN en una secuencia definida lo cual limitaría el riesgo de mutagénesis insercional.

En este trabajo se ha investigado el uso de meganucleasas y TALEN para estudiar facilitar la inserción específica de genes marcadores y terapéuticos, particularmente el gen *FANCA*, en una nueva región del genoma.

1. EXPERIMENTOS DE EDICIÓN GÉNICA EN CÉLULAS HEK-293H

La línea celular HEK-293H se caracteriza por su fácil manejo, su alto tasa de proliferación y su facilidad para ser transfectada utilizando, por ejemplo, Lipofectamina. Esto la convierte en una célula muy adecuada para testar la eficiencia de las nucleasas cedidas por nuestros colaboradores de Cellectis. Estas nucleasas, dos meganucleasas isoesquizómeros, SH6v5 y SH6v2, y una TALEN, tienen como diana determinadas secuencias dentro del sitio seguro determinado SH6. Para facilitar la edición génica se empleó un plásmido matriz, al que hemos denominado GM, que contenía el marcador EGFP controlado por el promotor EF1α largo, caracterizado por inducir una expresión estable. El promotor y el gen *EGFP* estaban flanqueados por secuencias homólogas a la secuencia del SH6.

En primer lugar testamos la funcionalidad del matriz GM en las células transfectando 4 µg de este plásmido frente a otro plásmido portador de los genes marcadores GFP/Cherry (**Figure 23**). En ambas condiciones se observó un alto porcentaje de fluorescentes, indicando una buena funcionalidad de ambas construcciones. Sin embargo, el porcentaje era superior en las células transfectadas con la matriz GM, por lo que un número de copias que habían sido

transfectadas por célula probablemente fuera mayor. El porcentaje de células $EGFP^+$ fue disminuyendo con el paso del tiempo hasta su completa desaparición, lo que podría deberse a la degradación del ADN no integrado en el genoma de la célula. Podemos concluir que nuestra matriz se transfecta fácilmente en las células, pero no se integra por sí mismo a pesar de llevar los brazos de homología.

Respecto a los experimentos de edición génica, comenzamos comparando las dos MNs diseñadas para el sitio SH6 y usando la GM como molde para la recombinación homóloga. En primer lugar analizamos si a medida que pasaba el tiempo, el pool de células transfectadas mantenía un porcentaje estable de células $EGFP^+$, lo cual sería una medida indirecta de la integración de la matriz en el genoma. Se comprobó que, tras 24 días en cultivo, el 4% de las células eran $EGFP^+$ en la condición en la que se usó la SH6v5-MN, cuatro veces más que en las demás condiciones, lo que podría ser indicativo de que las meganucleasas facilitan la integración específica de la GM en el sitio SH6.

Respecto a las eficiencias de las meganucleasas para inducir HR, se ha publicado un trabajo que refiere a una meganucleasa con especificidad para el sitio SH6. Eyquem *et al.* trabajaron con tres meganucleasas con especificidad para tres loci diferentes, entre ellos el SH6. Consiguieron eficiencias de entre un 1-2% para este sitio seguro trabajando con un molde consistente en un plásmido con un EGFP (flanqueado por brazos de homologías de 1 kb en lugar de 1,5kb) en células HEK-293H, siguiendo el mismo protocolo de transfección y análisis y usando dosis de MN más bajas o similares a las nuestras (0,03 y 0,3 µg) (Eyquem, Poirot et al. 2013). En nuestro caso conseguimos eficiencias de recombinación homóloga, con el isoesquizómero SH6v5-MN, de entre 4,5% y 15%. Teniendo en cuenta que tras la transfección de las células se siembran 10 células por pocillo, y estimando una frecuencia de clonaje del 30%, las eficiencias que obtendríamos serían de entre un 1,5% y un 5%, un poco mayores que las que ellos proponen.

Nuestros colaboradores de Cellectis también nos cedieron una TALEN con especificidad por el SH6, por lo que comparamos esta SH6-TALEN con la SH6v5-MN en HEK-293H para verificar cuál de las dos resultaba ser la más eficiente. En primer lugar comprobamos cuál de las dos era más activa utilizando el ensayo Surveyor con la dosis optimizada de la MN y varias dosis de la TAEN (**Figure 29**). Ambas nucleasas presentaron actividades similares en estas células, puesto que las dos bandas adicionales mostraban intensidades similares entre las distintas condiciones. En la dosis más alta de la TALEN se observó que la actividad de la nucleasa es algo menor que con dosis más bajas. Esto podría deberse al efecto citotóxico generado por las altas dosis de la TALEN. Posteriormente se analizó el porcentaje de células $EGFP^+$ en el total de células transfectadas, comprobándose a día 34 post-transfección que con las dosis de 0,5, 3,0 y 10 µg de TALEN el porcentaje de células expresando EGFP era 2 veces mayor que la condición transfectada sólo con la GM y 1,4 veces más que en las células transfectadas con la MN. Como se discutió anteriormente, esto es una medida indirecta de la integración de la GM mediada por las nucleasas, lo que sugeriría que la TALEN podría ser más eficiente respecto a las MNs.

A la hora de comparar de forma directa la eficiencia de cada nucleasa, se obtuvieron porcentajes similares tanto con la MN como con los 0,3, 3,0 y 10 µg de TALEN: de 15 al 20% (5-6,6% teniendo en cuenta el 30% de frecuencia de clonaje).

Cabe destacar que, en algunos experimentos en los que las transfecciones fueron mayores y la actividad de corte de la nucleasa fue mucho más alta, los porcentajes de HR bajaron de forma considerable. Esto vincula un corte muy alto de la nucleasa con la citotoxicidad y, por tanto, con la bajada de eficiencia de la recombinaión homóloga. En vista de los resultados obtenidos se eligieron para continuar trabajando dosis de 3 μg de TALEN. Por otra parte, para establecer comparaciones en la eficiencia de la TALEN frente a las MNS es necesario tener en cuenta que para una determinada cantidad de TALEN hay que poner el doble de ADN que con la MN, y de un plásmido más grande, pues la TALEN funciona como heterodímero.

Después de haber calculado la eficiencia aproximada de nuestras nucleasas con especificidad para el SH6, establecimos pseudoclones de los pocillos que contenían células editadas específicamente para llevar a cabo más estudios. Comprobamos la expresión de EGFP a tiempos más prolongados después de la transfección, para evaluar su estabilidad y la funcionalidad del transgén integrado en el sitio seguro SH6, así como que la integración realmente se había llevado a cabo de forma específica en el SH6 (**Figure 27** y **Figure 31**).

En algunos casos se observó que no todos los pseudoclones resultaron positivos a la PCR específica (**Figure 31**). Esto podría ser debido a la naturaleza heterogénea de los pseudoclones, y a la potencial citotoxicidad generada por las nucleasas. Así mismo comprobamos que no todos los pseudoclones mantenían la expresión de EGFP probablemente también debido a la naturaleza heterogénea de los pseudoclones.

También se observó que la intensidad de expresión del EGFP variaba de un pseudoclón a otro. Esto lo explicaría el número de copias que integradas en el SH6, (este número de copias se discutirá en el párrafo de la discusión de los resultados del *Southern*). La expresión de un trangén depende tanto del locus en el que se integre como del promotor que controla su expresión, de modo que si en estos pseudoclones hubiera más de una copia del transgén integrada, la intensidad de expresión del mismo aumentaría. En el trabajo de Eyquem *et al.*, compararon la eficiencia de MNs con dianas en tres genes distintos, entre ellos el SH6, y estudiaron la expresión del EGFP del donador en estos tres *loci* usando tres promotores diferentes, entre ellos el EF1α (que lleva un intrón) usado en nuestro trabajo. Comprobaron que la elección del promotor tenía una gran importancia sobre la expresión del transgén en un determinado locus. El promotor EF1α convencional (sin intrón) no permitió expresión del EGFP en ninguno de los *loci* estudiados para clones con una única copia pero los promotores SFFV y el EF1α + intrón (como el usado en nuestro trabajo) sí que permitieron que el transgén se expresara en los tres *loci*. Sin embargo, el nivel de esta expresión era diferente en función del locus, de modo que el EF1α + intrón mostró una menor intensidad en la expresión del transgén en los tres *loci*, sobretodo en el SH6, comparado al nivel conseguido por el promotor SFFV, aunque esta expresión era considerablemente más uniforme con el promotor EF1α + intrón (Eyquem, Poirot et al. 2013). La menor expresión del transgén observado en el *locus* SH6, se debe sobre todo a que se trata de una región no activa transcripcionalmente por tratarse precisamente de un desierto genético. Según la base de datos ENCODE, en esta región del genoma en 7 líneas celulares, la histona 3 está muy poco acetilada y su lisina 4 débilmente metilada. Este tipo de modificaciones están relacionadas con una cromatina poco permisiva para la expresión de un transgén (Yan and Boyd 2006, Eyquem, Poirot et al. 2013). Los bajos niveles de expresión en el SH6 no supondrían sin embargo un problema en el desarrollo de

protocolos de terapia génica con el gen *FANCA* de la anemia de Fanconi, puesto que se ha estudiado que bajos niveles de expresión de este gen son suficientes para revertir el fenotipo de células AF-A (Gonzalez-Murillo, Lozano et al. 2010).

Para comprobar la especificidad de la interación de la GM en el sitio SH6, realizamos ensayos de *Southern blot* usando como sonda el EGFP marcado radiactivamente con P^{32}. Los pseudoclones fueron digeridos con BamHI, que tiene una diana dentro de la matriz y otra en el genoma, y que en caso de integración específica debería dar una banda de 3.946 pb. En un primer *Southern*, [en el experimento de comparación de las SH6-MN (SH6v2 y SH6v5), mostrado en la **Figure 28**], se analizaron los 7 pseudoclones que fueron positivos para integración específica según la PCR. Seis de los 7 mostraron una banda del tamaño correcto (1, 2, 4, 5, 6 y 7), indicando integración específica en el sitio SH6. Sin embargo, 4 de estos pseudoclones mostraron una banda adicional más grande, de aproximadamente 7.242 pb. El pseudoclón 3 no mostró ninguna banda a pesar de ser positivo para la PCR. Sin embargo, este pseudoclón fue el que presentó un nivel de expresión de EGFP más bajo, de modo que la ausencia de banda podría deberse a la baja proporción de células editadas en este pseudoclón y que la técnica no fuera capaz de detectarlo (**Figure 27**). En el experimento de comparación entre la SH6v5-MN y la SH6-TALEN se lograron establecer varios pseudoclones, 7 de los cuales se analizaron por *Southern blot* como en el caso anterior (**Figure 32**). Esta vez, de los 7 pseudoclones, 3 presentaron tanto la banda esperada como la más grande (2, 3 y 7), los mismos pseudoclones que presentaron niveles altos o intermedios de expresión de EGFP y que fueron de nuevo positivos para la PCR específica. Sin embargo, uno de estos clones, el 6, que había sido positivo para la PCR específica, no presentó banda, coincidiendo de nuevo con el hecho de ser el pseudoclón con el menor porcentaje de células $EGFP^+$ (un 16,2%) (**Figure 31**).

Con respecto a la banda de 7.242 pb, ésta podría deberse a la formación de concatémeros del donador, en nuestro caso la GM, durante el proceso de integración mediada por HR. Este fenómeno ya ha sido observado en otros trabajos. Por ejemplo, Lombardo *et al.* nucleofectaron linfocitos T con ZFN y un IDLV donador y comprobaron al realizar el análisis de *Southern* que aparecían bandas adicionales de mayor tamaño al esperado (Lombardo, Cesana et al. 2011). También en el trabajo de Eyquem *et al.* en el que usaron MNs para tres *loci* diferentes, entre ellos el SH6, observaron este fenómeno (Eyquem, Poirot et al. 2013), que además iba acompañado de un mayor nivel de expresión del transgén, como se ha observado en nuestro trabajo.

Por último, realizamos un estudio comparativo de eficiencias de integración específica con nuestras nucleasas utilizando otro donador portador del gen *FANCA* junto con el gen de resistencia a puromicina. Mientras que el tamaño total de la matriz GM fue de 7.440 pb, el de la matriz terapéutica, la TM, era de 11,848 pb lo que sin duda podría limitar la eficacia delas nucleasas y de la transfección además de incrementar la citotoxicidad. Por este motivo comparamos la viabilidad de las células transfectadas con la matriz GM frente a las tranfectadas con la matriz TM solas, o junto a la MN o la TALEN (**Figure 34**). En estos experimentos se observó que no inducía una toxicidad importante por la combinación nucleasa + GM.

Cuando comparamos eficiencias de HR con las dos matrices, sólo obtuvimos eventos de HR con la matriz TM cuando la combinamos con la MN, siendo la eficiencia de HR la mitad a la obtenida usando la matriz GM (**Figure 35**). Esto es consistente con la evidencia de que donadores más grandes son menos eficientes a la hora de integrarse por HR en el genoma. Con la TALEN sólo se obtuvieron células editadas genéticamente cuando se combinó con la matriz GM, pero no con la terapéutica. Esto podría sugerir que la TALEN no es tan eficiente como la SH5v5-MN para inducir HR con el donador terapéutico.

Para comprobar la funcionalidad de la matriz terapéutica una vez integrada en el SH6, serían necesarios experimentos complementarios en los que se analizase la expresión de la proteína mediante Western blot con un anticuerpo anti-3flag (marcador que lleva el *FANCA* de la matiz) para testar la expresión de este transgén. No sería de utilidad un anticuerpo anti-*FANCA* porque esta línea celular ya lo posee porque no proviene de un paciente. La realización experimentos de edición génica en células obtenidas de pacientes, por ejemplo LCLs y células $CD34^+$, facilitaría demostrarla funcionalidad de la matriz terapéutica, pues su correcto funcionamiento revertiría el fenotipo enfermo y les conferiría resistencia frente a agentes generadores de entrecruzamientos en el ADN, tales como MMC.

2. EXPERIMENTOS DE EDICIÓN GÉNICA EN CÉLULAS PROGENITORAS HEMATOPOYÉTICAS $CD34^+$

Una vez que se comprobó en la línea celular HEK-293H que tanto la SH6v5-MN y la SH6-TALEN eran capaces de generar células editadas en las que el donador GM se insertaba específicamente en el sitio seguro SH6 con baja citotoxicidad, decidimos trabajar en prognitores hematopoyéticos $CD34^+$ de sangre de cordón .

La célula diana ideal para enfermedades monogénicas que afectan al sistema hematopoyético, y por supuesto, para pacientes con anemia de Fanconi serían las células progenitoras hematopoyéticas, como las $CD34^+$. En la introducción, concretamente en la **Table 2**, están resumidos todos los ensayos clínicos de terapia génica de enfermedades monogénicas del sistema hematopoyético llevados a cabo, incluidos en anemia de Fanconi, en algunos de los cuales se usaron estas células como diana. En concreto, la existencia de pacientes mosaico en esta enfermedad, en los que una mutación espontánea en una células madre hematopoyética es capaz de revertir el fenotipo de Fanconi, debido a que dota a estas células de una ventaja proliferativa que le permite crecer y diferenciarse hasta repoblar todo el sistema hematopoyético del paciente, convierte a estas células en la diana perfecta para llevar a cabo la terapia génica (Gross, Hanenberg et al. 2002, Antonio Casado, Callen et al. 2007). Además, teniendo en cuenta que la terapia génica dirigida minimiza el riesgo de producir oncogénesis insercional en las células tratadas y que ya existe un antecedente en el que fibroblastos de pacientes de Fanconi han sido editados usando ZFNs y un donador como IDLV (Rio, Banos et al. 2014), pensamos que ya había una base sólida que indicaba que la terapia génica con

nucleasas en células CD34$^+$ podía funcionar y tener, para la AF, perspectivas de aplicación terapéutica.

En primer lugar decidimos llevar a cabo una serie de análisis para comprobar la integridad de las CD34$^+$ durante el protocolo para llevar a cabo la edición génica. Las células madre hematopoyéticas son células delicadas cuyas propiedades pluripotentes pueden verse afectadas por factores externos, ya sean las condiciones de cultivo *in vitro,* como la citotoxicidad producida tanto por los métodos de transfección empleados o por la acción de las nucleasas o el ADN transfectado. El hecho de que sólo haya dos trabajos publicados en los que se hace edición génica por recombinación homóloga en precursores hematopoyéticos CD34$^+$ es una prueba más de la dificultad y el reto que entraña trabajar con estas células (Genovese, Schiroli et al. 2014, Hoban, Cost et al. 2015).

Uno de los primeros análisis que llevamos a cabo fue testar cuánto ADN podían soportar las células, nucleofectándolas con el *device I* del AMAXA, usando dosis crecientes de un plásmido con dos genes reporteros, *GFP* y *Cherry*. En general, las células soportaron sin problemas el efecto de la nucleofección y hasta 15 µg de este ADN, aunque a mayores dosis la toxicidad aumentó y los niveles de expresión transitoia de los genes marcadores iban disminuyendo (**Figure 37B**).Cuando nucleofectamos las células con las nucleasas y 4µg del donador GM, obtuvimos buenas viabilidades y mejores porcentajes de células que expreasban el transgén EGFP de nuestro donador, indicando la viabilidad de nuestra aproximación.

El mantenimiento de la capacidad para diferenciarse a los linajes mieloide y eritroide resulta de suma importancia en la función de las CMHs, pues en ella se basa la terapia génica de enfermedades monogénicas. Por este motivo, realizamos análisis para comprobar si las células que habían sido nucleofectadas con la matriz GM, y después sorteadas (CD34$^+$/EGFP$^+$), mantenían esta capacidad. Se observó que, aunque el número de colonias era menor que en células no manipuladas, se obtenían colonias diferenciadas en linajes hematopoyéticos mieloide y eritroide, lo que confirmaba el mantenimiento de su capacidad clonogénica a pesar de todo el proceso, tal y como se muestra en la **Figure 38**.

El uso del separador celular nos resultó interesante, en un principio, como medio para enriquecer las células nucleofectadas, pues las células que expresaban el EGFP varios días después de la nucleofección serían células que contenían el donador GM y que, por tanto, podrían ser las que con mayor probabilidad podrían integrarlo en el sitio seguro SH6. Se obtuvieron colonias de los dos linajes y algunas de ellas expresaban el EGFP. Estas colonias EGFP$^+$ se picaron y se analizaron mediante una reacción de PCR para identificar aquellas en las que el EGFP se había integrado de forma específica en el SH6. Sin embargo ninguna de estas colonias dio positivo a este studio. Esto nos llevó a pensar que tuvieran integraciones off-target que resultaran en una alta expresión del transgén. El SH6 es un sitio hipercompactado en el que no se espera que se produzcan estas expresiones tan altas de los genes integrados. Esto también está contemplado en el trabajo de Eyquem *et al*. (Eyquem, Poirot et al. 2013). También podría ocurrir que el transgén expresado de manera trans, y que, por tanto, al analizar las colonias el resultado fuera negativo.

Otro factor a tener en cuenta para el mantenimiento de la pluripotencialidad de las CD34$^+$es la combinación de citoquinas empleados para el mantenimiento de las células durante todo el

proceso de edición génica. Los estímulos agresivos que las $CD34^+$reciben, como la nucleofección o el cultivo *in vitro*, o la citotoxicidad generada por la entrada de ADN puede inducir la diferenciación de los progenitores. En los trabajos publicados sobre edición génica de células $CD34^+$ se han utilizado unas condiciones diferentes de cultivo. Hoban *et al.* utilizan ZFNs para el tratamiento SCD y en su cultivo usan medio X-VIVO 15 con 50 ng/mL de hSCF, hTPO y h Flt3-L y, aunque no realizan estudios de poblaciones, sí comprobaron que las células editadas se diferenciaban al linaje eritropoyético mediante cultivos de metilcelulosa, obteniendo un número de colonias similar al que nosotros hemos obtenido en nuestro proceso de edición génica (Hoban, Cost et al. 2015). Genovese *et al.* emplearon medio StemSpan con P/S y 100 ng/mL de hSCF y hFlt3-L junto con 20 ng/mL hTPO y hIL-6. Estos autores observaron que la adición al medio de cultivo de 1 µM de SR1 y 10 µM de dmPGE2 aumentaba la capacidad de autorrenovación de las $CD34^+$, así como de reconstituir la hematopoyesis a largo plazo en ratones NSG (Genovese, Schiroli et al. 2014). En nuestro caso hemos utilizado una combinación diferente de citoquinas consistente en StemSpan con P/S y 300 ng/mL de hSCF, hTPO y hFlt3-L. Como se muestra en la **Figure 49**, en los experimentos llevados a cabo se observó que hasta un 11% de la población nucleofectada pertenecía a la subpoblación de CMHs más primitiva, y hasta un 4% de la población total era además positiva para EGFP, indicando que al menos el donador GM se expresaba al menos de forma transitoria en células con comportamiento de CMHs.

La nucleofección suele ser el método de elección para introducir ADN y ARN en células madre hematopoyéticas si no es posible hacerlo con LVs. Aunque el tamaño del ADN es un factor importante para las dos técnicas, cuando la secuencia codificante de la nucleasa o del propio casete donador es superior a unas 5 kb, la producción de LV se ve seriamente afectada. Por este motivo, nucleofectar nucleasas y matrices mediante la electroporación de ADN plasmídico nos pareció eñ método de elección por el gran tamaño de las TALEN, y por el del gen terapéutico de *FANCA* que constituye nuestra matriz terapéutica. Sin embargo, el aspecto más negativo de la nucleofección radica en que el proceso daña a las células y también limita la cantidad de ADN que podemos introducir. En el experimento mostrado en la **Figure 50** y la **Figure 51** se consiguió editar células $CD34^+$ usando 5 µg de SH6-TALEN tal como indica la PCR específica en el conjunto de células nucleofectadas, así como en una de las colonias obtenidas después del ensayo clonogénico. El principal inconveniente de los resultados de este experimento fue la gran toxicidad observada en dicha condición, principalmente debido a la gran cantidad de ADN nucleofectado para introducir las dos subunidades de la TALEN, y el propio proceso de nucleofección. Esto nos hizo pensar en otras alternativas para intentar reducir el daño que nuestra metodología generaba a estas células, como por ejemplo sintetizar la TALEN como un mRNA y cambiar del nucleofector *Device I* al 4D de Lonza.

La nucleofección o transfección de las nucleasas como ARNm cada vez tiene más aplicación en la terapia génica dirigida. Su uso elimina el riesgo de la mutagénesis insercional y disminuye la probabilidad de que se produzcan off-targets debido a que tiene una vida media más corta que el ADN. Esto permite que el tiempo en el que la nucleasa es activa en la células es más bajo, disminuyendo así su citotoxicidad, (Maggio and Goncalves 2015, Skipper and Mikkelsen 2015). Sin embargo, sintetizar ARNm de nucleasas complejas y de mayor tamaño, como por ejemplo de las TALEN, es complicado, pues secuencias tan largas son inestables y se degradan con facilidad. Por este motivo clonamos el 3´UTR de la β-globina tras la secuencia que codifica la

TALEN, para sintetizar un ARNm más estable (Ji, Kong et al. 2011), Además, ya existen trabajos en los que se han empleado las TALEn como ARN para editar linfocitos T (Poirot, Philip et al. 2015). Después de clonar esta secuencia en los plásmidos que codificaban las subunidades de la SH6-TALEN con el promor T7, el ARNm se sintetizó y se comprobó la actividad de la nucleasa tras haber sido nucleofectada en las células $CD34^+$. Se comparó su actividad con otro ARNm sintetizado de plásmidos sin esta secuencia nucleofectando los progenitores y analizando la actividad de la nucleasa por el Surveyor assay (**Figure 52**). La nucleasa nucleofectada como ARNm con la secuencia 3´UTR era más activa en comparación con la misma nucleofectada como ADN o sin el 3´UTR de la β-globina. Además, la viabilidad observada después de la nucleofección era de hasta 3 veces superior en las condiciones tratadas con la nucleasa como ARNm que como ADN. Estos resultados sugieren que la región 3´UTR estabiliza el ARN, lo que aumentaría el tiempo de vida y actividad de la nucleasa, traduciéndose en una mayor cantidad de corte en comparación al ARN que no contiene esta secuencia. Sin embargo, en ninguno de los experimentos de edición génica llevados a cabo en en este trabajo en el que se ha usado la TALEN como ARNm en células $CD34^+$ se consiguió introducir el donador GM en el SH6. En Genovese *et al.* nucleofectaron células $CD34^+$ con 175 µg/mL de ZFN como ARNm utilizando como donador un IDLV consiguiendo HR en el gen *IL2RG*. En el estudio de Hoban *et al.* las células $CD34^+$ fueron nucleofectadas con dosis crecientes de ARNm (10-30 µg/mL) de ZFNs específicos para cortar a ambos lados de la mutación que origina la SCD en el gen de la β-globina. Finalmente trabajaron con dosis de 10 µg/mL de ARNm de ZFN, que les permitía un equilibrio entre la edición génica y la viabilidad celular. Si comparamos estos trabajos con el nuestro, comprobamos que en el caso del grupo de Genovese, las dosis de ARNm empleadas son mayores que las nuestras, pero no en el caso del trabajo de Hoban, de modo que la dosis empleada no tendría por qué ser un problema. Además, si comparamos los nucleofectores, nosotros hemos empleado el mismo equipo que Genovese con resultados negativos. Esto nos lleva a pensar que la diferencia estriba en el gen que queremos editar, pues nosotros trabajamos en el SH6 que, como comentamos anteriormente en la discusión de las 293H, se trata de una región hipermetilada (Eyquem, Poirot et al. 2013), lo que podría dificultar el acceso de la nucleasa a esta región. Las células progenitoras fueron editadas únicamente al utilizarse la SH6-TALEN como ADN a dosis más altas. Esto permitiría que la nucleasa fuera expresada durante más tiempo en la célula y. En esto hay que tener en cuenta que las células progenitoras se caracterizan por no tener un ratio de división y proliferación muy alto, hecho que se considera una dificultad para editar por HR estas células por no pasar muchas veces por la fase S del ciclo celular, donde es más probable que la HR tenga lugar y donde además la cromatina está menos compactada, lo que facilitaría el acceso de la nucleasa al SH6.

Respecto al nucleofector empleado, nosotros sólo conseguimos editar las células $CD34^+$ con el Device I de AMAXA, usando para ello una dosis más alta de SH6-TALEN como DNA y sin sortear las células que expresaban EGFP. Debido a la toxicidad que se generaba, decidimos probar el nucleofector 4D, caracterizado por ser menos agresivo que el anterior. Testamos tanto las cubetas como los strips, nucleofectando 10^6 células $CD34^+$ en las cubetas y 10^5 con los strips. Como estas células son difíciles de obtener, decidimos probar los strips para facilitar el trabajo con las mismas. Sin embargo, tal y como se muestra en la **Figure 53**, con las cubetas obteníamos unos porcentajes de eficiencia de nucleofección del molde GM del 45-60%, similar a los obtenido con el Device I, y con viabilidades similares. Se realizaron varios experimentos

con esta plataforma, incluso incrementando las dosis de nucleasa (tanto como ADN como ARNm), pero en ningún experimento con el nucleofector 4D observamos colonias editadas de forma específica con la matriz GM. Esto indicaba que, con este nucleofector, la eficiencia de las nucleasas era menor aun manteniendo todas las condiciones del protocolo iguales. Algo que observamos durante el desarrollo de este trabajo es que, entre todos los experimentos realizados, las eficiencias de nucleofección eran más reproducibles con el Device I que con el 4D (entre el 45-60%), dato que hemos comprobado es imprescindible como buen pronóstico para el éxito del proceso de edición génica. Aquellos experimentos en los que este porcentaje era menor no se observaba eventos de HR. En vista a los resultados obtenidos decidimos utilizar el Device I, observando eventos de edición génica tanto en el conjunto de las células nucleofectadas como en colonias individuales en las condiciones en las que se usaron dosis de 5 y 10 µg de ADN total de la SH6-TALEN.

Una explicación que aclare por qué el nucleofector 4D no funciona con nuestro protocolo podría ser por la metodología que utilizamos para introducir las nucleasa y el molde. El grupo de Genovese *et al*. nucleofecta la nucleasa como ARNm, pero como molde usa un IDLV, mientras que nosotros un plásmido que nucleofectamos junto con las nucleasas. Esto podría estar reduciendo la eficacia de las nucleasas para inducir HR.

Con respecto a las eficiencias de edición génica en el SH6, hay 3 puntos a tener en cuenta: la PCR específica en el pool de células, en las colonias, y con qué nucleasa tiene lugar la edición génica. Se analizaron en el pool de células la presencia de edición génica en el SH6 con el donor GM como control de funcionalidad del experimento. Una ausencia de la misma indicaría que sería muy poco probable observar integración específica en el SH6 en las colonias. Sin embargo, su presencia no implica que también se observen en las colonias, pues estas células con integración específica en el SH6 podrían formar parte de la población más diferenciada, de modo que en los ensayos de metilcelulosa no generarían colonias y no se observarían resultados positivos en las mismas. Esto explicaría también que en los análisis de actividad de las nucleasa se observe corte en las células en el sitio diana pero no eventos de edición génica.

Además, sólo se han observado colonias con integración específica con la matriz GM en las condiciones nucleofectadas con dosis altas de la SH6-TALEN, mientras que en los experimentos llevados a cabo en HEK-293H se observaron porcentajes similares de eficiencia para las dos nucleasas, pero no en $CD34^+$. Esto resulta curioso puesto que, teniendo en cuenta que el SH6 es un desierto genético fuertemente compactado (Eyquem, Poirot et al. 2013), cabría esperar que una nucleasa más pequeña, como esta meganucleasa, tuviera mejor acceso a la diana que una TALEN, que es más grande. Lo que sí está claro es que cómo esté la cromatina afecta al acceso de la MN, de la matriz de reparación (en nuestro caso la GM) y de la maquinaria de reparación al sitio diana, y está aceptado que el contexto cromosómico y la epigenética afectan la edición génica inducida por las MNs (Daboussi, Zaslavskiy et al. 2012). También se ha publicado un trabajo en el que se ha visto que metilaciones en dinucleótidos CpG en las dianas de las MNs en el genoma afecta a su afinidad y a su eficiencia (Valton, Daboussi et al. 2012). El conjunto de estas circunstancias podrían ser las causantes de que no se hayan conseguido editar las células $CD34^+$ con nuestro donador GM.

Respecto a los porcentajes de edición génica observados en este trabajo, pensamos que no son comparables con los otros dos trabajos publicados en $CD34^+$. En estos trabajos los porcentajes de edición génica son superiores a los que nosotros hemos conseguido. Sin embargo es importante tener en cuenta que no sólo utilizan una nucleasa diferente (ZFNs), sino que la aproximación también es diferente, puesto que ambos utilizan como molde un IDLV y en Hoban *et al,.* además, un oligonucleótido. Pero quizá la diferencia más importante radica en que ellos trabajan en genes distintos al utilizado en nuestro trabajo que además permiten una mejor expresión, mientras que el SH6 es un desierto genético. Las diferencias de eficiencia de las nucleasas varían mucho en función del gene diana con el que se trabaja.

3. OTROS ASPECTOS A TENER EN CUENTA

Uno de los puntos importantes a la hora de llevar a cabo la edición génica es la metodología que hemos de emplear para conseguir que tanto las nucleasas como el molde que se integrará en el sitio diana por recombinación homóloga, entren de forma eficiente dentro de la célula. IDLVs, plásmidos y ARNm son las plataformas más utilizadas para conseguir este objetivo. Sin embargo, el ARNm sintetizado *in vitro* de las nucleasas es cada vez más usado para la mayoría de estas aplicaciones. Que tenga una menor vida media dentro de las células posibilita que la acción de las nucleasas sea más limitada en el tiempo, disminuyendo la citotoxicidad en las células y la posibilidad de off-targets. Aunque en este trabajo se ha utilizado esta aproximación, también se han empleado ADN plásmídico para la introducción de las nucleasas. El principal motivo es que las TALEN son muy grandes y, además de que sería complicado obtener un título aceptable a la hora de producir LVs, se ha visto que es muy frecuente la recombinación entre las secuencias repetitivas del dominio de unión al ADN de las TALEN durante la producción de los mismos (Holkers, Maggio et al. 2013). Su tamaño también sería un problema para sintetizarla como ARNm, pero añadiendo la secuencia 3´UTR de la β-globina se consiguiría estabilizalo. Sin embargo, en este trabajo no se han observado eventos de edición génica utilizando la SH6-TALEN como ARNm, aunque su actividad era incluso más alta en el sitio diana que la observada con la introducida como ADN plasmídico. El hecho de que el SH6 sea una región del genoma hipercompactada, junto con la vida media menor de la TALEN como ARNm podría ser determinante para estos resultados. También cabe la posibilidad de que se necesite una mayor cantidad de ARNm, pues en otros trabajos las dosis utilizadas eran mayores (Genovese, Schiroli et al. 2014).

Con respecto al donador, lo más utilizado siguen siendo los IDLV, aunque se ha visto que su uso puede estar asociado con eventos de recombinación ilegítima adicional (Holkers, Maggio et al. 2014). En este trabajo se utilizaron plásmidos debido a la necesidad de un soporte que pudiera contener una gran cantidad de ADN, pues el gen *FANCA* con el que se pretende la aproximación terapéutica en AF es muy grande (4.365 pb) y, junto con el promotor, el gen de resistencia o reportero y los brazos de homología alcanzaría un tamaño de 9.518 pb. Tanto con

el donador GM como con el TM se observaró integración específica en las HEK-293H, y con la GM en las células $CD34^+$, demostrando la funcionalidad de este soporte.

Trabajar con el SH6 entraña una cierta dificultad debido a que se trata de un desierto genético que está hipercompactado, y más metilado que otras regiones que están más abiertas porque están transcribiéndose. Esto supondría un problema para el acceso de las nucleasas al sitio diana, así como a la maquinaria de reparación del ADN. Sin embargo, en las células HEK-293H se ha conseguido la integración específica de ambos donadores con ambas nucleasas con una expresión estable. En las células $CD34^+$ se consiguió empleando la SH6-TALEN como ADN y la matriz GM, aunque no se observó la expresión del transgén EGFP en las colonias positivas. Sin embargo, esta baja expresión no tiene por qué ser un problema, puesto que está estudiado que bajos niveles de expresión de la proteína FANCA son suficientes para revertir el fenotipo de la enfermedad (Gonzalez-Murillo, Lozano et al. 2010).

Una posible alternativa para mitigar la influencia de esta cromatina tan compactada sería el uso de agentes demetilantes como el AZA antes de la nucleofección. Tal y como comentábamos antes, cuando los dinucleótidos CpG en las dianas de las MNs en el genoma están metilados, afecta a su afinidad y a su eficiencia y observaron que la adición del AZA su eficiencia aumentaba (Valton, Daboussi et al. 2012).

Otro punto que también cabe destacar es que el objetivo de este trabajo, tal como su título indica, era el desarrollo de nuevas alternativas terapéuticas para tratar la anemia de Fanconi. La terapia génica dirigida se basa en la utilización de la maquinaria interna que la célula usa para reparar su ADN para introducir un gen reportero o terapéutico exógenos por recombinación homóloga, en este caso en el SH6. Esto es importante en una enfermedad que está caracterizada por tener deteriorada su maquinaria de reparación del ADN. Aunque no se tiene demasiado claro cuál es exactamente la función de cada uno de los genes implicados en la ruta, excepto *BRCA2*, cuya ausencia claramente afecta a la recombinación homóloga, sí que se sabe que mutaciones en los genes que forman partes del Core-complex, como *FANCA*, tienen un efecto más leve, permitiendo esta estrategia terapéutica (Xia, Taghian et al. 2001, Nakanishi, Cavallo et al. 2011). Además, existen trabajos publicado en los que se ha conseguido hacer edición génica en células de Fanconi, como se describe en el punto 3.3 de la Introducción: “Gene editing and Fanconi anemia”.

Más experimentos tienen que llevarse a cabo con el protocolo diseñado en este trabajo. En células $CD34^+$ sería interesante intentar comprobar si hay off-target después del proceso de edición génica. Debido a la limitación que supone el número de células, una posible estrategia podría ser generar iPS desde un número de células nucleofectadas, de forma que pudieran crecer indefinidamente y obtener un número de células sufuciente para hacer pruebas como Southern blot. También, como se comentó en la discusión de las HEK-293H, sería interesante probar nuestro donador terapéutico en células obtenidas de pacientes, como LCLs, o interferir células $CD34^+$ para simular en las mismas el fenotipo de Fanconi.

CONCLUSIONS

VIII. CONCLUSIONS

1. We have constructed a Green Matrix donor carrying the *EGFP* reporter gene driven by the EF1α promoter, flanked by homology arms to the SH6 safe harbor *locus*.

2. We have demonstrated the specific integration of the Green Matrix in the SH6 site of HEK293-H cells using specific meganucleases and TALEN. In some instances, however, concatemeric insertions of the Green Matrix in the SH6 were observed.

3. *In vitro* studies showed that the integration of the Green Matrix in the SH6 site of HEK293-H cells allows the stable expression of the *EGFP* reporter gene.

4. We have constructed a Therapeutic Matrix, in which the *FANCA* and the puromycin-resistant genes replaced the *EGFP* reporter gene in the Green Matrix.

5. We have demonstrated the specific integration of the Therapeutic Matrix in the SH6 site of HEK293 cells using the SH6v5-meganuclease.

6. Studies conducted in primary $CD34^+$ cells from human cord blood have shown that, under our optimized conditions, 3.1% of the human hematopoietic progenitor cells harbor specific integrations of the Green Matrix in the *SH6 locus*.

Taken together, our results suggest that the SH6 site may constitute a relevant new safe harbor *locus* for conducting targeted gene therapy in human hematopoietic progenitors and stem cells. Our observations would have a particular impact in diseases such as Fanconi anemia, in which the proliferation advantage of a reduced number of corrected stem cells may have a significant clinical impact.

Further studies aiming to demonstrate the stable expression of therapeutic genes inserted in the SH6 *locus* would be required, however, prior to consider our targeted gene therapy approach for clinical applications.

BIBLIOGRAPHY

IX. BIBLIOGRAPHY

Adam, R., D. Azoulay, D. Castaing, R. Eshkenazy, G. Pascal, K. Hashizume, D. Samuel and H. Bismuth (2003). "Liver resection as a bridge to transplantation for hepatocellular carcinoma on cirrhosis: a reasonable strategy?" Ann Surg 238(4): 508-518; discussion 518-509.

Aiuti, A., L. Biasco, S. Scaramuzza, F. Ferrua, M. P. Cicalese, C. Baricordi, F. Dionisio, A. Calabria, S. Giannelli, M. C. Castiello, M. Bosticardo, C. Evangelio, A. Assanelli, M. Casiraghi, S. Di Nunzio, L. Callegaro, C. Benati, P. Rizzardi, D. Pellin, C. Di Serio, M. Schmidt, C. Von Kalle, J. Gardner, N. Mehta, V. Neduva, D. J. Dow, A. Galy, R. Miniero, A. Finocchi, A. Metin, P. P. Banerjee, J. S. Orange, S. Galimberti, M. G. Valsecchi, A. Biffi, E. Montini, A. Villa, F. Ciceri, M. G. Roncarolo and L. Naldini (2013). "Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome." Science 341(6148): 1233151.

Aiuti, A., F. Cattaneo, S. Galimberti, U. Benninghoff, B. Cassani, L. Callegaro, S. Scaramuzza, G. Andolfi, M. Mirolo, I. Brigida, A. Tabucchi, F. Carlucci, M. Eibl, M. Aker, S. Slavin, H. Al-Mousa, A. Al Ghonaium, A. Ferster, A. Duppenthaler, L. Notarangelo, U. Wintergerst, R. H. Buckley, M. Bregni, S. Marktel, M. G. Valsecchi, P. Rossi, F. Ciceri, R. Miniero, C. Bordignon and M. G. Roncarolo (2009). "Gene therapy for immunodeficiency due to adenosine deaminase deficiency." N Engl J Med 360(5): 447-458.

Annoni, A., K. Goudy, M. Akbarpour, L. Naldini and M. G. Roncarolo (2013). "Immune responses in liver-directed lentiviral gene therapy." Transl Res 161(4): 230-240.

Antonio Casado, J., E. Callen, A. Jacome, P. Rio, M. Castella, S. Lobitz, T. Ferro, A. Munoz, J. Sevilla, A. Cantalejo, E. Cela, J. Cervera, J. Sanchez-Calero, I. Badell, J. Estella, A. Dasi, T. Olive, J. Jose Ortega, A. Rodriguez-Villa, M. Tapia, A. Molines, L. Madero, J. C. Segovia, K. Neveling, R. Kalb, D. Schindler, H. Hanenberg, J. Surralles and J. A. Bueren (2007). "A comprehensive strategy for the subtyping of patients with Fanconi anaemia: conclusions from the Spanish Fanconi Anemia Research Network." J Med Genet 44(4): 241-249.

Arnould, S., P. Chames, C. Perez, E. Lacroix, A. Duclert, J. C. Epinat, F. Stricher, A. S. Petit, A. Patin, S. Guillier, S. Rolland, J. Prieto, F. J. Blanco, J. Bravo, G. Montoya, L. Serrano, P. Duchateau and F. Paques (2006). "Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets." J Mol Biol 355(3): 443-458.

Auerbach, A. D., A. Rogatko and T. M. Schroeder-Kurth (1989). "International Fanconi Anemia Registry: relation of clinical symptoms to diepoxybutane sensitivity." Blood 73(2): 391-396.

Bacman, S. R., S. L. Williams, M. Pinto, S. Peralta and C. T. Moraes (2013). "Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs." Nat Med 19(9): 1111-1113.

Bejerano, G., M. Pheasant, I. Makunin, S. Stephen, K. W. James, J. S. Mattick and D. Haussler (2004). "Ultraconserved Elements in the Human Genome." Science 304: 5.

Bessis, N., F. J. GarciaCozar and M. C. Boissier (2004). "Immune responses to gene therapy vectors: influence on vector function and effector mechanisms." Gene Ther 11 Suppl 1: S10-17.

Bhatia, M., D. Bonnet, B. Murdoch, O. I. Gan and J. E. Dick (1998). "A newly discovered class of human hematopoietic cells with SCID-repopulating activity." Nat Med 4(9): 1038-1045.

Bianchi, M., A. Hakkim, V. Brinkmann, U. Siler, R. A. Seger, A. Zychlinsky and J. Reichenbach (2009). "Restoration of NET formation by gene therapy in CGD controls aspergillosis." Blood 114(13): 2619-2622.

Bianchi, M., M. J. Niemiec, U. Siler, C. F. Urban and J. Reichenbach (2011). "Restoration of anti-Aspergillus defense by neutrophil extracellular traps in human chronic granulomatous disease after gene therapy is calprotectin-dependent." J Allergy Clin Immunol 127(5): 1243-1252 e1247.

Bibikova, M., K. Beumer, J. K. Trautman and D. Carroll (2003). "Enhancing gene targeting with designed zinc finger nucleases." Science 300(5620): 764.

Bibikova, M., D. Carroll, D. J. Segal, J. K. Trautman, J. Smith, Y. G. Kim and S. Chandrasegaran (2001). "Stimulation of homologous recombination through targeted cleavage by chimeric nucleases." Mol Cell Biol 21(1): 289-297.

Bibikova, M., M. Golic, K. G. Golic and D. Carroll (2002). "Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases." Genetics 161(3): 1169-1175.

Bielorai, B., M. R. Hughes, A. D. Auerbach, A. Nagler, R. Loewenthal, G. Rechavi and A. Toren (2004). "Successful umbilical cord blood transplantation for Fanconi anemia using preimplantation genetic diagnosis for HLA-matched donor." Am J Hematol 77(4): 397-399.

Biffi, A., E. Montini, L. Lorioli, M. Cesani, F. Fumagalli, T. Plati, C. Baldoli, S. Martino, A. Calabria, S. Canale, F. Benedicenti, G. Vallanti, L. Biasco, S. Leo, N. Kabbara, G. Zanetti, W. B. Rizzo, N. A. Mehta, M. P. Cicalese, M. Casiraghi, J. J. Boelens, U. Del Carro, D. J. Dow, M. Schmidt, A. Assanelli, V. Neduva, C. Di Serio, E. Stupka, J. Gardner, C. von Kalle, C. Bordignon, F. Ciceri, A. Rovelli, M. G. Roncarolo, A. Aiuti, M. Sessa and L. Naldini (2013). "Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy." Science 341(6148): 1233158.

Bloom, K., A. Ely, C. Mussolino, T. Cathomen and P. Arbuthnot (2013). "Inactivation of hepatitis B virus replication in cultured cells and in vivo with engineered transcription activator-like effector nucleases." Mol Ther 21(10): 1889-1897.

Boch, J. and U. Bonas (2010). "Xanthomonas AvrBs3 family-type III effectors: discovery and function." Annu Rev Phytopathol 48: 419-436.

Boch, J., H. Scholze, S. Schornack, A. Landgraf, S. Hahn, S. Kay, T. Lahaye, A. Nickstadt and U. Bonas (2009). "Breaking the code of DNA binding specificity of TAL-type III effectors." Science 326(5959): 1509-1512.

Bogdanove, A. J., S. Schornack and T. Lahaye (2010). "TAL effectors: finding plant genes for disease and defense." Curr Opin Plant Biol 13(4): 394-401.

Bogdanove, A. J. and D. F. Voytas (2011). "TAL effectors: customizable proteins for DNA targeting." Science 333(6051): 1843-1846.

Bogliolo, M., B. Schuster, C. Stoepker, B. Derkunt, Y. Su, A. Raams, J. P. Trujillo, J. Minguillon, M. J. Ramirez, R. Pujol, J. A. Casado, R. Banos, P. Rio, K. Knies, S. Zuniga, J. Benitez, J. A. Bueren, N. G. Jaspers, O. D. Scharer, J. P. de Winter, D. Schindler and J. Surralles (2013). "Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia." Am J Hum Genet 92(5): 800-806.

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

Bonfim, C. M., C. R. de Medeiros, M. A. Bitencourt, J. Zanis-Neto, V. A. Funke, D. C. Setubal, J. Ruiz, J. E. Sanders, M. E. Flowers, H. P. Kiem, R. Storb and R. Pasquini (2007). "HLA-matched related donor hematopoietic cell transplantation in 43 patients with Fanconi anemia conditioned with 60 mg/kg of cyclophosphamide." Biol Blood Marrow Transplant 13(12): 1455-1460.

Boztug, K., M. Schmidt, A. Schwarzer, P. P. Banerjee, I. A. Diez, R. A. Dewey, M. Bohm, A. Nowrouzi, C. R. Ball, H. Glimm, S. Naundorf, K. Kuhlcke, R. Blasczyk, I. Kondratenko, L. Marodi, J. S. Orange, C. von Kalle and C. Klein (2010). "Stem-cell gene therapy for the Wiskott-Aldrich syndrome." N Engl J Med 363(20): 1918-1927.

Braun, C. J., K. Boztug, A. Paruzynski, M. Witzel, A. Schwarzer, M. Rothe, U. Modlich, R. Beier, G. Gohring, D. Steinemann, R. Fronza, C. R. Ball, R. Haemmerle, S. Naundorf, K. Kuhlcke, M. Rose, C. Fraser, L. Mathias, R. Ferrari, M. R. Abboud, W. Al-Herz, I. Kondratenko, L. Marodi, H. Glimm, B. Schlegelberger, A. Schambach, M. H. Albert, M. Schmidt, C. von Kalle and C. Klein (2014). "Gene therapy for Wiskott-Aldrich syndrome--long-term efficacy and genotoxicity." Sci Transl Med 6(227): 227ra233.

Callen, E., J. A. Casado, M. D. Tischkowitz, J. A. Bueren, A. Creus, R. Marcos, A. Dasi, J. M. Estella, A. Munoz, J. J. Ortega, J. de Winter, H. Joenje, D. Schindler, H. Hanenberg, S. V. Hodgson, C. G. Mathew and J. Surralles (2005). "A common founder mutation in FANCA underlies the world's highest prevalence of Fanconi anemia in Gypsy families from Spain." Blood 105(5): 1946-1949.

Candotti, F., K. L. Shaw, L. Muul, D. Carbonaro, R. Sokolic, C. Choi, S. H. Schurman, E. Garabedian, C. Kesserwan, G. J. Jagadeesh, P. Y. Fu, E. Gschweng, A. Cooper, J. F. Tisdale, K. I. Weinberg, G. M. Crooks, N. Kapoor, A. Shah, H. Abdel-Azim, X. J. Yu, M. Smogorzewska, A. S. Wayne, H. M. Rosenblatt, C. M. Davis, C. Hanson, R. G. Rishi, X. Wang, D. Gjertson, O. O. Yang, A. Balamurugan, G. Bauer, J. A. Ireland, B. C. Engel, G. M. Podsakoff, M. S. Hershfield, R. M. Blaese, R. Parkman and D. B. Kohn (2012). "Gene therapy for adenosine deaminase-deficient severe combined immune deficiency: clinical comparison of retroviral vectors and treatment plans." Blood 120(18): 3635-3646.

Capecchi, M. R. (1989). "Altering the genome by homologous recombination." Science 244(4910): 1288-1292.

Carlson, D. F., W. Tan, S. G. Lillico, D. Stverakova, C. Proudfoot, M. Christian, D. F. Voytas, C. R. Long, C. B. Whitelaw and S. C. Fahrenkrug (2012). "Efficient TALEN-mediated gene knockout in livestock." Proc Natl Acad Sci U S A 109(43): 17382-17387.

Carroll, D. (2011). "Zinc-finger nucleases: a panoramic view." Curr Gene Ther 11(1): 2-10.

Cartier, N., S. Hacein-Bey-Abina, C. C. Bartholomae, G. Veres, M. Schmidt, I. Kutschera, M. Vidaud, U. Abel, L. Dal-Cortivo, L. Caccavelli, N. Mahlaoui, V. Kiermer, D. Mittelstaedt, C. Bellesme, N. Lahlou, F. Lefrere, S. Blanche, M. Audit, E. Payen, P. Leboulch, B. l'Homme, P. Bougneres, C. Von Kalle, A. Fischer, M. Cavazzana-Calvo and P. Aubourg (2009). "Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy." Science 326(5954): 818-823.

Castella, M., R. Pujol, E. Callen, M. J. Ramirez, J. A. Casado, M. Talavera, T. Ferro, A. Munoz, J. Sevilla, L. Madero, E. Cela, C. Belendez, C. D. de Heredia, T. Olive, J. S. de Toledo, I. Badell, J. Estella, A. Dasi, A. Rodriguez-Villa, P. Gomez, M. Tapia, A. Molines, A. Figuera, J. A. Bueren and J. Surralles (2011). "Chromosome fragility in patients with Fanconi anaemia: diagnostic implications and clinical impact." J Med Genet 48(4): 242-250.

Castella, M., R. Pujol, E. Callen, J. P. Trujillo, J. A. Casado, H. Gille, F. P. Lach, A. D. Auerbach, D. Schindler, J. Benitez, B. Porto, T. Ferro, A. Munoz, J. Sevilla, L. Madero, E. Cela, C. Belendez, C. D. de Heredia, T. Olive, J. S. de Toledo, I. Badell, M. Torrent, J. Estella, A. Dasi, A. Rodriguez-Villa, P. Gomez, J. Barbot, M. Tapia, A. Molines, A. Figuera, J. A. Bueren and J. Surralles (2011). "Origin, functional role, and clinical impact of Fanconi anemia FANCA mutations." Blood 117(14): 3759-3769.

Castiello, M. C., S. Scaramuzza, F. Pala, F. Ferrua, P. Uva, I. Brigida, L. Sereni, M. van der Burg, G. Ottaviano, M. H. Albert, M. Grazia Roncarolo, L. Naldini, A. Aiuti, A. Villa and M. Bosticardo (2015). "B-cell reconstitution after lentiviral vector-mediated gene therapy in patients with Wiskott-Aldrich syndrome." J Allergy Clin Immunol 136(3): 692-702 e692.

Cathomen, T. and J. K. Joung (2008). "Zinc-finger nucleases: the next generation emerges." Mol Ther 16(7): 1200-1207.

Cavazzana-Calvo, M., E. Payen, O. Negre, G. Wang, K. Hehir, F. Fusil, J. Down, M. Denaro, T. Brady, K. Westerman, R. Cavallesco, B. Gillet-Legrand, L. Caccavelli, R. Sgarra, L. Maouche-Chretien, F. Bernaudin, R. Girot, R. Dorazio, G. J. Mulder, A. Polack, A. Bank, J. Soulier, J. Larghero, N. Kabbara, B. Dalle, B. Gourmel, G. Socie, S. Chretien, N. Cartier, P. Aubourg, A. Fischer, K. Cornetta, F. Galacteros, Y. Beuzard, E. Gluckman, F. Bushman, S. Hacein-Bey-Abina and P. Leboulch (2010). "Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia." Nature 467(7313): 318-322.

Ciccia, A. and S. J. Elledge (2010). "The DNA damage response: making it safe to play with knives." Mol Cell 40(2): 179-204.

Ciccia, A., C. Ling, R. Coulthard, Z. Yan, Y. Xue, A. R. Meetei, H. Laghmani el, H. Joenje, N. McDonald, J. P. de Winter, W. Wang and S. C. West (2007). "Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM." Mol Cell 25(3): 331-343.

Civin, C. I., T. Trischmann, N. S. Kadan, J. Davis, S. Noga, K. Cohen, B. Duffy, I. Groenewegen, J. Wiley, P. Law, A. Hardwick, F. Oldham and A. Gee (1996). "Highly purified CD34-positive cells reconstitute hematopoiesis." J Clin Oncol 14(8): 2224-2233.

Collis, S. J., A. Ciccia, A. J. Deans, Z. Horejsi, J. S. Martin, S. L. Maslen, J. M. Skehel, S. J. Elledge, S. C. West and S. J. Boulton (2008). "FANCM and FAAP24 function in ATR-mediated checkpoint signaling independently of the Fanconi anemia core complex." Mol Cell 32(3): 313-324.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini and F. Zhang (2013). "Multiplex genome engineering using CRISPR/Cas systems." Science 339(6121): 819-823.

Crossan, G. P. and K. J. Patel (2012). "The Fanconi anaemia pathway orchestrates incisions at sites of crosslinked DNA." J Pathol 226(2): 326-337.

Chandrasegaran, S. and J. Smith (1999). "Chimeric restriction enzymes: what is next?" Biol Chem 380(7-8): 841-848.

Chandrasekharappa, S. C., F. P. Lach, D. C. Kimble, A. Kamat, J. K. Teer, F. X. Donovan, E. Flynn, S. K. Sen, S. Thongthip, E. Sanborn, A. Smogorzewska, A. D. Auerbach, E. A. Ostrander and N. C. S. Program (2013). "Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia." Blood 121(22): e138-148.

Chang, D. J. and K. A. Cimprich (2009). "DNA damage tolerance: when it's OK to make mistakes." Nat Chem Biol 5(2): 82-90.

Chen, J., W. Zhang, J. Lin, F. Wang, M. Wu, C. Chen, Y. Zheng, X. Peng, J. Li and Z. Yuan (2014). "An efficient antiviral strategy for targeting hepatitis B virus genome using transcription activator-like effector nucleases." Mol Ther 22(2): 303-311.

Chen, Z. and H. Zhao (2005). "A highly sensitive selection method for directed evolution of homing endonucleases." Nucleic Acids Res 33(18): e154.

Cheng, A. W., H. Wang, H. Yang, L. Shi, Y. Katz, T. W. Theunissen, S. Rangarajan, C. S. Shivalila, D. B. Dadon and R. Jaenisch (2013). "Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system." Cell Res 23(10): 1163-1171.

Chinen, J., J. Davis, S. S. De Ravin, B. N. Hay, A. P. Hsu, G. F. Linton, N. Naumann, E. Y. Nomicos, C. Silvin, J. Ulrick, N. L. Whiting-Theobald, H. L. Malech and J. M. Puck (2007). "Gene therapy improves immune function in preadolescents with X-linked severe combined immunodeficiency." Blood 110(1): 67-73.

Cho, S. W., S. Kim, J. M. Kim and J. S. Kim (2013). "Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease." Nat Biotechnol 31(3): 230-232.

Cho, S. W., S. Kim, Y. Kim, J. Kweon, H. S. Kim, S. Bae and J. S. Kim (2014). "Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases." Genome Res 24(1): 132-141.

Chylinski, K., A. Le Rhun and E. Charpentier (2013). "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems." RNA Biol 10(5): 726-737.

D'Andrea, A. D. and M. Grompe (1997). "Molecular biology of Fanconi anemia: implications for diagnosis and therapy." Blood 90(5): 1725-1736.

Daboussi, F., M. Zaslavskiy, L. Poirot, M. Loperfido, A. Gouble, V. Guyot, S. Leduc, R. Galetto, S. Grizot, D. Oficjalska, C. Perez, F. Delacote, A. Dupuy, I. Chion-Sotinel, D. Le Clerre, C. Lebuhotel, O. Danos, F. Lemaire, K. Oussedik, F. Cedrone, J. C. Epinat, J. Smith, R. J. Yanez-Munoz, G. Dickson, L. Popplewell, T. Koo, T. VandenDriessche, M. K. Chuah, A. Duclert, P. Duchateau and F. Paques (2012). "Chromosomal context and epigenetic mechanisms control the efficacy of genome editing by rare-cutting designer endonucleases." Nucleic Acids Res 40(13): 6367-6379.

De Palma, M., E. Montini, F. R. Santoni de Sio, F. Benedicenti, A. Gentile, E. Medico and L. Naldini (2005). "Promoter trapping reveals significant differences in integration site selection between MLV and HIV vectors in primary hematopoietic cells." Blood 105(6): 2307-2315.

de Winter, J. P., F. Leveille, C. G. van Berkel, M. A. Rooimans, L. van Der Weel, J. Steltenpool, I. Demuth, N. V. Morgan, N. Alon, L. Bosnoyan-Collins, J. Lightfoot, P. A. Leegwater, Q. Waisfisz, K. Komatsu, F. Arwert, J. C. Pronk, C. G. Mathew, M. Digweed, M. Buchwald and H. Joenje (2000). "Isolation of a cDNA representing the Fanconi anemia complementation group E gene." Am J Hum Genet 67(5): 1306-1308.

de Winter, J. P., M. A. Rooimans, L. van Der Weel, C. G. van Berkel, N. Alon, L. Bosnoyan-Collins, J. de Groot, Y. Zhi, Q. Waisfisz, J. C. Pronk, F. Arwert, C. G. Mathew, R. J. Scheper, M. E. Hoatlin, M. Buchwald and H. Joenje (2000). "The Fanconi anaemia gene FANCF encodes a novel protein with homology to ROM." Nat Genet 24(1): 15-16.

de Winter, J. P., Q. Waisfisz, M. A. Rooimans, C. G. van Berkel, L. Bosnoyan-Collins, N. Alon, M. Carreau, O. Bender, I. Demuth, D. Schindler, J. C. Pronk, F. Arwert, H. Hoehn, M. Digweed, M. Buchwald and H. Joenje (1998). "The Fanconi anaemia group G gene FANCG is identical with XRCC9." Nat Genet 20(3): 281-283.

Deeg, H. J., G. Socie, G. Schoch, M. Henry-Amar, R. P. Witherspoon, A. Devergie, K. M. Sullivan, E. Gluckman and R. Storb (1996). "Malignancies after marrow transplantation for aplastic anemia and fanconi anemia: a joint Seattle and Paris analysis of results in 700 patients." Blood 87(1): 386-392.

Delacote, F., C. Perez, V. Guyot, M. Duhamel, C. Rochon, N. Ollivier, R. Macmaster, G. H. Silva, F. Paques, F. Daboussi and P. Duchateau (2013). "High frequency targeted mutagenesis using engineered endonucleases and DNA-end processing enzymes." PLoS One 8(1): e53217.

Deng, D., C. Yan, X. Pan, M. Mahfouz, J. Wang, J. K. Zhu, Y. Shi and N. Yan (2012). "Structural basis for sequence-specific recognition of DNA by TAL effectors." Science 335(6069): 720-723.

Ding, Q., Y. K. Lee, E. A. Schaefer, D. T. Peters, A. Veres, K. Kim, N. Kuperwasser, D. L. Motola, T. B. Meissner, W. T. Hendriks, M. Trevisan, R. M. Gupta, A. Moisan, E. Banks, M. Friesen, R. T. Schinzel, F. Xia, A. Tang, Y. Xia, E. Figueroa, A. Wann, T. Ahfeldt, L. Daheron, F. Zhang, L. L. Rubin, L. F. Peng, R. T. Chung, K. Musunuru and C. A. Cowan (2013). "A TALEN genome-editing system for generating human stem cell-based disease models." Cell Stem Cell 12(2): 238-251.

Ding, Q., A. Strong, K. M. Patel, S. L. Ng, B. S. Gosis, S. N. Regan, C. A. Cowan, D. J. Rader and K. Musunuru (2014). "Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing." Circ Res 115(5): 488-492.

Dorsman, J. C., M. Levitus, D. Rockx, M. A. Rooimans, A. B. Oostra, A. Haitjema, S. T. Bakker, J. Steltenpool, D. Schuler, S. Mohan, D. Schindler, F. Arwert, G. Pals, C. G. Mathew, Q. Waisfisz, J. P. de Winter and H. Joenje (2007). "Identification of the Fanconi anemia complementation group I gene, FANCI." Cell Oncol 29(3): 211-218.

Doulatov, S., F. Notta, E. Laurenti and J. E. Dick (2012). "Hematopoiesis: a human perspective." Cell Stem Cell 10(2): 120-136.

Doyon, J. B., V. Pattanayak, C. B. Meyer and D. R. Liu (2006). "Directed evolution and substrate specificity profile of homing endonuclease I-SceI." J Am Chem Soc 128(7): 2477-2484.

Du, W., O. Erden and Q. Pang (2014). "TNF-alpha signaling in Fanconi anemia." Blood Cells Mol Dis 52(1): 2-11.

Duan, X., F. S. Gimble and F. A. Quiocho (1997). "Crystal structure of PI-SceI, a homing endonuclease with protein splicing activity." Cell 89(4): 555-564.

Dufour, C. and J. Svahn (2008). "Fanconi anaemia: new strategies." Bone Marrow Transplant 41 Suppl 2: S90-95.

Ellis, J. (2005). "Silencing and variegation of gammaretrovirus and lentivirus vectors." Hum Gene Ther 16(11): 1241-1246.

Epinat, J. C., S. Arnould, P. Chames, P. Rochaix, D. Desfontaines, C. Puzin, A. Patin, A. Zanghellini, F. Paques and E. Lacroix (2003). "A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells." Nucleic Acids Res 31(11): 2952-2962.

Eyquem, J., L. Poirot and R. Galetto (2013). "Characterization of three loci for homologous gene targeting and transgene expression." Biotechnology Bioengineering 110: 11.

Fadel, H. J., J. H. Morrison, D. T. Saenz, J. R. Fuchs, M. Kvaratskhelia, S. C. Ekker and E. M. Poeschla (2014). "TALEN knockout of the PSIP1 gene in human cells: analyses of HIV-1 replication and allosteric integrase inhibitor mechanism." J Virol 88(17): 9704-9717.

Fanconi anaemia/Breast cancer, c. (1996). "Positional cloning of the Fanconi anaemia group A gene." Nat Genet 14(3): 324-328.

Flowers, G. P., A. T. Timberlake, K. C. McLean, J. R. Monaghan and C. M. Crews (2014). "Highly efficient targeted mutagenesis in axolotl using Cas9 RNA-guided nuclease." Development 141(10): 2165-2171.

Frank, S., B. V. Skryabin and B. Greber (2013). "A modified TALEN-based system for robust generation of knock-out human pluripotent stem cell lines and disease models." BMC Genomics 14: 773.

Friedland, A. E., Y. B. Tzur, K. M. Esvelt, M. P. Colaiacovo, G. M. Church and J. A. Calarco (2013). "Heritable genome editing in C. elegans via a CRISPR-Cas9 system." Nat Methods 10(8): 741-743.

Fu, Y., J. A. Foden, C. Khayter, M. L. Maeder, D. Reyon, J. K. Joung and J. D. Sander (2013). "High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells." Nat Biotechnol 31(9): 822-826.

Gabriel, R., A. Lombardo, A. Arens, J. C. Miller, P. Genovese, C. Kaeppel, A. Nowrouzi, C. C. Bartholomae, J. Wang, G. Friedman, M. C. Holmes, P. D. Gregory, H. Glimm, M. Schmidt, L. Naldini and C. von Kalle (2011). "An unbiased genome-wide analysis of zinc-finger nuclease specificity." Nat Biotechnol 29(9): 816-823.

Garcia-Higuera, I., Y. Kuang, J. Denham and A. D. D'Andrea (2000). "The fanconi anemia proteins FANCA and FANCG stabilize each other and promote the nuclear accumulation of the Fanconi anemia complex." Blood 96(9): 3224-3230.

Gasiunas, G. and V. Siksnys (2013). "RNA-dependent DNA endonuclease Cas9 of the CRISPR system: Holy Grail of genome editing?" Trends Microbiol 21(11): 562-567.

Gaspar, H. B., S. Cooray, K. C. Gilmour, K. L. Parsley, S. Adams, S. J. Howe, A. Al Ghonaium, J. Bayford, L. Brown, E. G. Davies, C. Kinnon and A. J. Thrasher (2011). "Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency." Sci Transl Med 3(97): 97ra79.

Gaspar, H. B., S. Cooray, K. C. Gilmour, K. L. Parsley, F. Zhang, S. Adams, E. Bjorkegren, J. Bayford, L. Brown, E. G. Davies, P. Veys, L. Fairbanks, V. Bordon, T. Petropoulou, C. Kinnon and A. J. Thrasher (2011). "Hematopoietic stem cell gene therapy for adenosine deaminase-deficient severe combined immunodeficiency leads to long-term immunological recovery and metabolic correction." Sci Transl Med 3(97): 97ra80.

Geibler, R., H. Scholze, S. Hahn, J. Streube, U. Bonas, S.-E. Behrens and J. Boch (2011). "Transcriptional Activators of Human Genes with Programmable DNA-Specificity." PLoS One 6(5).

Genovese, P., G. Schiroli, G. Escobar, T. D. Tomaso, C. Firrito, A. Calabria, D. Moi, R. Mazzieri, C. Bonini, M. C. Holmes, P. D. Gregory, M. v. d. Burg, B. Gentner, E. Montini, A. Lombardo and L. Naldini (2014). "Targeted genome editing in human repopulating haematopoietic stem cells." Nature 510: 6.

Ghosh, S., A. J. Thrasher and H. B. Gaspar (2015). "Gene therapy for monogenic disorders of the bone marrow." Br J Haematol.

Gilbert, L. A., M. H. Larson, L. Morsut, Z. Liu, G. A. Brar, S. E. Torres, N. Stern-Ginossar, O. Brandman, E. H. Whitehead, J. A. Doudna, W. A. Lim, J. S. Weissman and L. S. Qi (2013). "CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes." Cell 154(2): 442-451.

Goebel, W. S. and M. C. Dinauer (2003). "Gene therapy for chronic granulomatous disease." Acta Haematol 110(2-3): 86-92.

Gonzalez-Murillo, A., M. L. Lozano, L. Alvarez, A. Jacome, E. Almarza, S. Navarro, J. C. Segovia, H. Hanenberg, G. Guenechea, J. A. Bueren and P. Rio (2010). "Development of lentiviral vectors with optimized transcriptional activity for the gene therapy of patients with Fanconi anemia." Hum Gene Ther 21(5): 623-630.

Gratz, S. J., A. M. Cummings, J. N. Nguyen, D. C. Hamm, L. K. Donohue, M. M. Harrison, J. Wildonger and K. M. O'Connor-Giles (2013). "Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease." Genetics 194(4): 1029-1035.

Grewal, S. S., J. P. Kahn, M. L. MacMillan, N. K. Ramsay and J. E. Wagner (2004). "Successful hematopoietic stem cell transplantation for Fanconi anemia from an unaffected HLA-genotype-identical sibling selected using preimplantation genetic diagnosis." Blood 103(3): 1147-1151.

Grizot, S., J. Smith, F. Daboussi, J. Prieto, P. Redondo, N. Merino, M. Villate, S. Thomas, L. Lemaire, G. Montoya, F. J. Blanco, F. Paques and P. Duchateau (2009). "Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease." Nucleic Acids Res 37(16): 5405-5419.

Gross, M., H. Hanenberg, S. Lobitz, R. Friedl, S. Herterich, R. Dietrich, B. Gruhn, D. Schindler and H. Hoehn (2002). "Reverse mosaicism in Fanconi anemia: natural gene therapy via molecular self-correction." Cytogenet Genome Res 98(2-3): 126-135.

Guardiola, P., G. Socie, X. Li, P. Ribaud, A. Devergie, H. Esperou, P. Richard, R. Traineau, A. Janin and E. Gluckman (2004). "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 103(1): 73-77.

Guenechea, G., O. I. Gan, C. Dorrell and J. E. Dick (2001). "Distinct classes of human stem cells that differ in proliferative and self-renewal potential." Nat Immunol 2(1): 75-82.

Hacein-Bey-Abina, S., A. Garrigue, G. P. Wang, J. Soulier, A. Lim, E. Morillon, E. Clappier, L. Caccavelli, E. Delabesse, K. Beldjord, V. Asnafi, E. MacIntyre, L. Dal Cortivo, I. Radford, N. Brousse, F. Sigaux, D. Moshous, J. Hauer, A. Borkhardt, B. H. Belohradsky, U. Wintergerst, M. C. Velez, L. Leiva, R. Sorensen, N. Wulffraat, S. Blanche, F. D. Bushman, A. Fischer and M. Cavazzana-Calvo (2008). "Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1." J Clin Invest 118(9): 3132-3142.

Hacein-Bey-Abina, S., J. Hauer, A. Lim, C. Picard, G. P. Wang, C. C. Berry, C. Martinache, F. Rieux-Laucat, S. Latour, B. H. Belohradsky, L. Leiva, R. Sorensen, M. Debre, J. L. Casanova, S. Blanche, A. Durandy, F. D. Bushman, A. Fischer and M. Cavazzana-Calvo (2010). "Efficacy of gene therapy for X-linked severe combined immunodeficiency." N Engl J Med 363(4): 355-364.

Hacein-Bey-Abina, S., F. Le Deist, F. Carlier, C. Bouneaud, C. Hue, J. P. De Villartay, A. J. Thrasher, N. Wulffraat, R. Sorensen, S. Dupuis-Girod, A. Fischer, E. G. Davies, W. Kuis, L. Leiva and M. Cavazzana-Calvo (2002). "Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy." N Engl J Med 346(16): 1185-1193.

Hacein-Bey-Abina, S., S. Y. Pai, H. B. Gaspar, M. Armant, C. C. Berry, S. Blanche, J. Bleesing, J. Blondeau, H. de Boer, K. F. Buckland, L. Caccavelli, G. Cros, S. De Oliveira, K. S. Fernandez, D. Guo, C. E. Harris, G. Hopkins, L. E. Lehmann, A. Lim, W. B. London, J. C. van der Loo, N. Malani, F. Male, P. Malik, M. A. Marinovic, A. M. McNicol, D. Moshous, B. Neven, M. Oleastro, C. Picard, J. Ritz, C. Rivat, A. Schambach, K. L. Shaw, E. A. Sherman, L. E. Silberstein, E. Six, F. Touzot, A. Tsytsykova, J. Xu-Bayford, C. Baum, F. D. Bushman, A. Fischer, D. B. Kohn, A. H. Filipovich, L. D. Notarangelo, M. Cavazzana, D. A. Williams and A. J. Thrasher (2014). "A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency." N Engl J Med 371(15): 1407-1417.

Hacein-Bey Abina, S., H. B. Gaspar, J. Blondeau, L. Caccavelli, S. Charrier, K. Buckland, C. Picard, E. Six, N. Himoudi, K. Gilmour, A. M. McNicol, H. Hara, J. Xu-Bayford, C. Rivat, F. Touzot, F. Mavilio, A. Lim, J. M. Treluyer, S. Heritier, F. Lefrere, J. Magalon, I. Pengue-Koyi, G. Honnet, S. Blanche, E. A. Sherman, F. Male, C. Berry, N. Malani, F. D. Bushman, A. Fischer, A. J. Thrasher, A. Galy and M. Cavazzana (2015). "Outcomes following gene therapy in patients with severe Wiskott-Aldrich syndrome." JAMA 313(15): 1550-1563.

Heath, P. J., K. M. Stephens, R. J. Monnat, Jr. and B. L. Stoddard (1997). "The structure of I-Crel, a group I intron-encoded homing endonuclease." Nat Struct Biol 4(6): 468-476.

Hilton, I. B., A. M. D'Ippolito, C. M. Vockley, P. I. Thakore, G. E. Crawford, T. E. Reddy and C. A. Gersbach (2015). "Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers." Nat Biotechnol 33(5): 510-517.

Hira, A., K. Yoshida, K. Sato, Y. Okuno, Y. Shiraishi, K. Chiba, H. Tanaka, S. Miyano, A. Shimamoto, H. Tahara, E. Ito, S. Kojima, H. Kurumizaka, S. Ogawa, M. Takata, H. Yabe and M. Yabe (2015). "Mutations in the gene encoding the E2 conjugating enzyme UBE2T cause Fanconi anemia." Am J Hum Genet 96(6): 1001-1007.

Hoban, M. D., G. J. Cost, M. C. Mendel, Z. Romero, M. L. Kaufman, A. V. Joglekar, M. Ho, D. Lumaquin, D. Gray, G. R. Lill, A. R. Cooper, F. Urbinati, S. Senadheera, A. Zhu, P. Q. Liu, D. E. Paschon, L. Zhang, E. J. Rebar, A. Wilber, X. Wang, P. D. Gregory, M. C. Holmes, A. Reik, R. P. Hollis and D. B. Kohn (2015). "Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells." Blood 125(17): 2597-2604.

Holkers, M., I. Maggio, S. F. Henriques, J. M. Janssen, T. Cathomen and M. A. Goncalves (2014). "Adenoviral vector DNA for accurate genome editing with engineered nucleases." Nat Methods 11(10): 1051-1057.

Holkers, M., I. Maggio, J. Liu, J. M. Janssen, F. Miselli, C. Mussolino, A. Recchia, T. Cathomen and M. A. Goncalves (2013). "Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells." Nucleic Acids Res 41(5): e63.

Howe, S. J., M. R. Mansour, K. Schwarzwaelder, C. Bartholomae, M. Hubank, H. Kempski, M. H. Brugman, K. Pike-Overzet, S. J. Chatters, D. de Ridder, K. C. Gilmour, S. Adams, S. I. Thornhill, K. L. Parsley, F. J. Staal, R. E. Gale, D. C. Linch, J. Bayford, L. Brown, M. Quaye, C. Kinnon, P. Ancliff, D. K. Webb, M. Schmidt, C. von Kalle, H. B. Gaspar and A. J. Thrasher (2008). "Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients." J Clin Invest 118(9): 3143-3150.

Howlett, N. G., T. Taniguchi, S. G. Durkin, A. D. D'Andrea and T. W. Glover (2005). "The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability." Hum Mol Genet 14(5): 693-701.

Howlett, N. G., T. Taniguchi, S. Olson, B. Cox, Q. Waisfisz, C. De Die-Smulders, N. Persky, M. Grompe, H. Joenje, G. Pals, H. Ikeda, E. A. Fox and A. D. D'Andrea (2002). "Biallelic inactivation of BRCA2 in Fanconi anemia." Science 297(5581): 606-609.

Hsu, P. D., D. A. Scott, J. A. Weinstein, F. A. Ran, S. Konermann, V. Agarwala, Y. Li, E. J. Fine, X. Wu, O. Shalem, T. J. Cradick, L. A. Marraffini, G. Bao and F. Zhang (2013). "DNA targeting specificity of RNA-guided Cas9 nucleases." Nat Biotechnol 31(9): 827-832.

Hu, X., N. Chang, X. Wang, F. Zhou, X. Zhou, X. Zhu and J. W. Xiong (2013). "Heritable gene-targeting with gRNA/Cas9 in rats." Cell Res 23(11): 1322-1325.

Hu, Z., W. Ding, D. Zhu, L. Yu, X. Jiang, X. Wang, C. Zhang, L. Wang, T. Ji, D. Liu, D. He, X. Xia, T. Zhu, J. Wei, P. Wu, C. Wang, L. Xi, Q. Gao, G. Chen, R. Liu, K. Li, S. Li, S. Wang, J. Zhou, D. Ma and H. Wang (2015). "TALEN-mediated targeting of HPV oncogenes ameliorates HPV-related cervical malignancy." J Clin Invest 125(1): 425-436.

Huang, M., R. Kennedy, A. M. Ali, L. A. Moreau, A. R. Meetei, A. D. D'Andrea and C. C. Chen (2011). "Human MutS and FANCM complexes function as redundant DNA damage sensors in the Fanconi Anemia pathway." DNA Repair (Amst) 10(12): 1203-1212.

Hwang, W. Y., Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai, J. D. Sander, R. T. Peterson, J. R. Yeh and J. K. Joung (2013). "Efficient genome editing in zebrafish using a CRISPR-Cas system." Nat Biotechnol 31(3): 227-229.

International Human Genome Sequencing, C., E. Adekoya, M. Ait-Zahra, N. Allen, M. Anderson, S. Anderson, F. Anufriev, J. Ambruster, K. Ayele, J. Baker, J. Baldwin, N. Barna, V. Bastien, S. Batzoglou, R. Beckerly, F. Beda, J. Bernard, B. Birren, B. Blumensteil, L. Boguslavsky, B. Boukghalter, A. Brown, G. Burkett, J. Camarata, A. Campopiano, H. Carneiro, Z. Chen, Y. Choephal, M. Colangelo, S. Collins, A. Collymore, P. Cooke, C. Davis, T. Dawoe, K. DeArellano, K. Devon, K. Dewar, J. S. Diaz, S. Dodge, E. Donelan, K. Dorjee, M. Doyle, A. Dube, A. Dupes, M. Endrizzi, A. Farina, S. Faro, D. Ferguson, P. Ferriera, H. Fischer, W. Fitzhugh, K. Flaherty, K. Foley, R. Funke, D. Gage, J. Galagan, S. Gardyna, D. Gilbert, S. Ginde, A. Gomes, M. Goyette, J. Graham, L. Graham, E. Grandbois, N. Grand-Pierre, G. Grant, D. Gregoire, R. Guerrero, B. Hagos, K. Harris, D. Hart, B. Hatcher, A. Heaford, L. Horton, C. Hosage-Norman, J. Howland, B. Hulme, I. Iliev, R. Johnson, C. Jones, M. Joseph, M. Judd, L. Kann, A. Karatas, D. Kelley, M. Kelly, D. Lama, J. Lamazares, E. S. Lander, T. Landers, A. Lane, K. LaRocque, H. LeBlanc, J. P. Leger, J. Lehoczky, R. LeVine, D. Lewis, T. Lewis, C. Lieu, L. Linton, G. Liu, X. Liu, K. Locke, Y. Lokyitsang, P. Macdonald, R. Martinez, K. Maru, M. McCarthy, P. McEwan, T. McGhee, B. McGing, A. McGurk, K. McKernan, J. McLaughlin, R. McPheeters, L. Meneus, J. Mesirov, T. Mihova, J. Meldrim, L. Meneus, J. Mesirov, T. Mihova, C. Miranda, V. Mlenga, M. Modeski, G. Montello, W. Morris, J. Morrow, L. Mulrain, T. Murphy, J. Mychaleckyj, J. Naylor, C. Newes, T. Ngodup, C. Nguyen, T. Nguyen, C. D. Norbu, N. Norbu, C. Nusbaum, T. O'Connor, P. O'Donnell, Y. Okaf, D. O'Neil, J. O'Shea, S. Osman, M. Paresi, B. Pavlin, K. M. Peterson, P. Phunkang, N. Pierre, V. Pollara, C. Raymond, M. Rieback, B. Riley, C. Rise, P. Rogov, J. Roman, M. Roman, M. Rosetti, D. Rothman, A. Roy, K. Roycraft, R. Santos, S. Schauer, R. Schupbach, S. Seaman, A. Sheridan, C. Smith, C. Sougnez, T. Speece, B. Spencer, N. Stange-Thomann, N. Stojanovic, C. Stone, N. Strauss, A. Subramanian, J. Talamas, P. Tchuinga, M. Temelko, P. Tenzin, S. Tesfaye, J. Theodore, A. Tirrell, I. Torruella-Miller, T. Trac, M. Travers, N. Travis, J. Trigilio, E. Tsao, H. Vassiliev, R. Veil, A. Vo, A. Wagner, J. Walsh, T. Wangdi, J. Wierzbowski, B. Wilson, X. Wu, Z. Abdellah, A. Ahamdi, S. Ahmed, M. Aimable, R. Ainscough, J. Almeida, A. Ambler, K. Ambrose, K. Ambrose, D. Andrews, N. Andrews, H. Arbery, B. Archer, G. Ash, K. Ashcroft, J. Ashurst, R. Ashwell, D. Atkin, A. Atkinson, J. Atwood, K. Aubin, T. Avis, A. Babbage, J. Bacon, C. Bagguley, J. Bailey, A. Baker, S. Bardill, D. Barker, K. Barlow, L. Baron, A. Barrett, R. Bartlett, D. Basham, V. Basham, A. Bateman, K. Bates, C. Baynes, L. Beard, L. Beard, S. Beard, D. Beare, A. Beasley, O. Beasley, S. Beck, E. Bell, D. Bellerby, T. Bellerby, R. Bemrose, J. Bennett, D. Bentley, M. Berks, M. Berks, G. Bethel, C. Bird, E. Birney, H. Bissell, S. Blackburne-Maze, S. Blakey, R. Bonnett, R. Bonnett, R. Border, N. Brady, J. Bray, S. Bray-Allen, A. Bridgeman, J. Brook, S. Brooking, A. Brown, C. Brown, J. Brown, M. Brown, M. Brown, R. Bruskiewich, J. Bryant, D. Buck, V. Buckle, C. Budd, J. Burberry, D. Burford, J. Burgess, W. Burrill, C. Burrows, J. Burton, P. Butcher, A. Butler, M. Cairns, B. Caning, C. Carder, P. Carder, N. Carter, T. Cavanna, K. Chan, J. Chapman, R. Charles, T. Clothia, C. Chui, M. Clamp, A. Clark, G. Clark, G. Clark, K. Clark, S. Clark, S. Clark, B. Clarke, E. Clarke, K. Clarke, C. Clee, S. Clegg, K. Clifford, J. Coates, V. Cobley, A. Coffey, P. Coggill, C. Lotte, R. Collier, S. Collings, J. Collins, P. Collins, P. Connor, J. Conquer, D. Conroy, D. Constance, L. Cook, J. Cooper, R. Cooper, R. Cooper, T. Copsey, N. Corby, L. Cornell, R. Cornell, A. Cottage, A. Coulson, G. Coville, A. Cox, T. Cox, R. Coxhill, M. Craig, T. Crane, M. Crawley, V. Crew, J. Cuff, K. Culley, A. Cummings, K. Cummings, P. Cummings, A. Curran, V. Curwen, J. Cutts, R. Daniels, L. Davidson, J. Davies, J. Davies, N. Davies, R. Davies, J. Davies, E. Dawson, R. Deadman, P. Dean, S. Dear, F. Dearden, M. Delgado, P. Deloukas, J. Dennis, P. Dhami, C. Dibling, R. Dobbs, R. Dobson, C. Dockree, D. Doddington, S. Dodsworth, N. Doggett, A. Dunham, I. Dunham, A. Dunn, M. Dunn, R. Durbin, J. Durham, R. Dwyer, M. Earthrowl, T. Eastham, C. Edwards, K. Edwards, A. Ellington, M. Ellwood, B. Emberson, H. Errington, G. Evans, J. Evans, K. Evans, R. Evans, T. Feltwell, S. Fennell, R. Finn, T. Flack, K. Fleming, J. Flint, M. Flint, Y. Floyd, S. Footman, J. Fowler, D. Frame, M. Francis, S. Francis, J. Frankland, A. Fraser, D. Fraser, L. French, D.

Frost, J. Frost, L. Frost, C. Frost, L. Fuller, K. Fullerton, A. Gardner, P. Garner, J. Garnett, L. Gatland, L. Gatland, J. Ghori, B. Gibbs, D. Gibson, J. Gilbert, L. Gilby, C. Gilson, M. Gorton, D. Grafham, M. Grant, S. Grant, I. Gray, L. Green, J. Greenhalgh, J. Greenhill, P. Gregg, S. Gregory, E. Griffiths, M. Griffiths, I. Guthrie, R. Gwilliam, R. Hall, K. Halls, G. Hall-Tamlyn, J. Hamlett, S. Hammond, J. Hancock, A. Harding, J. Harley, D. Harper, G. Harper, G. Harradence, C. L. Harrison, R. Harrison, D. Hassan, N. Hawkins, K. Hawley, K. Hayes, P. Heath, R. Heathcott, C. Hembry, T. Herd, S. Hewitt, D. Higgs, G. Hillyard, R. Hinkins, S. J. Ho, D. Hodgson, M. Hoffs, J. Holden, J. Holdgate, E. Holloway, I. Holmes, S. Holmes, S. Holroyd, A. Hoooper, L. Hopewell, B. Hopkins, G. Hornett, G. Hornsby, T. Hornsby, S. Horsley, R. Horton, P. Howard, P. Howden, G. Howell, T. Hubbard, E. Huckle, J. Hughes, J. Hughes, L. Hull, H. Hummerich, S. Humphray, M. Humphries, A. Hunt, P. Hunt, S. Hunt, D. Hyde, M. Ince, J. Isherwood, J. Izatt, M. Izmajlowicz, N. Jareborg, B. Jassal, G. Jeffery, K. Jeffery, C. Jeffery, K. Jekosch, L. Jenkins, T. Johansen, C. Johnson, C. Johnson, D. Johnson, K. Jolley, A. Jones, C. Jones, J. Jones, M. Jones, M. Jones, S. Jones, S. Joseph, A. Joy, L. Joy, V. Joy, G. Joyce, M. Jubb, K. Karunaratne, M. Kay, D. Kaye, L. Kearney, S. Kelley, J. Kershaw, R. Kettleborough, C. Kidd, P. Kierstan, A. Kimberley, A. King, S. Kingsley, G. Klingle, A. Knights, A. Krogh, P. Laidlaw, M. Laing, G. Laird, C. Lambart, R. Lamble, R. Lamble, C. Langford, T. Lau, S. Lawlor, S. Leather, M. Lehvaslaiho, S. Leonard, D. Leongamornlert, M. Leversha, J. Lightning, S. Lindsay, M. Line, S. Linsdell, P. Little, C. Lloyd, D. Lloyd, V. Lock, W. Lock, A. Lodziak, I. Longden, H. Loraine, R. Lord, J. Lovell, G. Lye, N. Marriott, A. Marrone, P. Marsden, V. Marsh, M. Martin, S. Martin, G. Maslen, D. Mason, L. Matthews, P. Matthews, M. Maynard, O. McCann, J. McClay, C. McCollum, L. McConnachie, B. McDonald, L. McDonald, J. McDowall, C. McKeown, S. McLauren, K. McLay, J. McLean, J. McMurdo, A. McMurray, D. McMurray, N. McWilliams, N. Mehta, N. Menuge, S. Mercer, A. Miah, G. Micklem, S. Miles, S. Milne, D. Mistry, S. Mistry, J. Mitchell, J. Mitchell, M. Mohammadi, C. Molina, P. Mooney, M. Moore, A. Moreland, B. Mortimore, R. Mott, J. Mullikan, B. Munday, E. Munday, A. Mungall, C. Murname, K. Murrell, A. Myers, D. Negus, D. Niblett, J. Nicholson, T. Nickerson, S. Nijjar, Z. Ning, J. Nisbet, C. Odell, D. O'Donovan, F. Ogbighele, T. Oinn, H. Oliver, K. Oliver, H. Orbell, A. Osborn, J. Osborne, E. Overton-Larty, C. Parkin, K. Parkin, G. Parry-Brown, D. Patel, R. Patel, A. Pearce, D. Pearson, A. Peck, R. Peck, J. Peden, C. Percy, A. Perito, I. Perrault, A. Peters, R. Pettett, B. Phillimore, K. Phillips, S. Phillips, D. Platt, E. Playford, B. Plumb, M. Pocock, K. Porter, C. Potter, S. Potter, D. Powell, R. Prathalingham, M. Quail, C. Quince, M. Quareshi, H. Ramsay, Y. Ramsey, S. Ranby, R. Rance, V. Rand, J. Ratford, L. Ratford, D. Read, D. Redhead, C. Rees, M. Reid, A. Reinhardt, A. Rice, C. Rice, P. Rice, S. Richard, S. Richardson, K. Ridler, L. Riethoven, M. Robinson, R. Rochford, J. Rogers, L. Rogers, H. Ross, M. Ross, A. Rule, J. Rule, B. Russell, J. Rutter, K. Safdar, N. Salter, J. Santoyo-Lopez, D. Saunders, C. Scott, D. Scott, I. Scott, F. Seager, M. Searle, P. Searle, H. Sehra, J. Shardelow, G. Sharp, T. Shaw, C. Shaw-Smith, J. Shearing, K. Sheppard, K. Sheppard, R. Sheppard, E. Sheridan, R. Shownkeen, R. Silk, M. Sims, S. Sims, S. Sivadasan, C. Skuce, L. Smink, A. Smith, L. Smith, L. Smith, M. Smith, R. Smith, S. Smith, H. Sneath, C. Soderlund, V. Solovyev, E. Sonhammer, E. Sotheran, L. Spraggon, J. Squares, S. Squares, M. Stables, J. Stalker, S. Stamford, M. Stammers, H. Steingruber, Y. Stephens, C. Steward, A. Stewart, M. Stewart, M. Stock, L. Stoppard, P. Storey, J. Sturdy, J. Sulston, C. Swainson, M. Swann, N. Sycamore, M. Tagney, S. Tan, E. Tarling, A. Taylor, G. Taylor, K. Tauylor, R. Taylor, S. Taylor, L. Tee, J. Tester, A. Theaker, C. Thomas, D. Thomas, K. Thomas, R. Thomas, R. Thommai, A. Thorpe, K. Thorpe, G. Threadgold, E. Tinsley, A. Tracey, J. Travers, A. Tromans, B. Tubby, C. Tufarelli, K. Turney, D. Upson, M. Vaudin, R. Viknaraja, W. Vine, P. Voak, S. Walker, M. Wall, J. Wallis, M. Wallis, G. Warren, G. Warry, A. Watson, A. Webb, J. Webb, A. Wells, S. Wells, R. Welton, P. West, T. West, A. Wheatley, C. Wheatley, G. Wheeler, H. Whitaker, A. White, A. White, B. White, J. White, S. White, M. Whiteley, A. Whittaker, P. Whittaker, S. Widaa, A. Wild, J. Wilkinson, P. Wilkinson, D. Willey, A.

Williams, B. Williams, L. Williams, S. Williams, H. Williamson, T. Wilmer, L. Wilming, B. Wilson, G. Wilson, M. Wilson, N. Wilson, S. Wilson, W. Wilson, P. Window, J. Winster, J. Witt, F. Wobus, E. Wood, J. Wood, S. Woodeson, R. Woodhouse, R. Wooster, M. Wray, P. Wray, K. Wright, J. Wyatt, J. Xie, L. Young, S. Young, R. Younger, S. Zhao, S. Abbas, A. Abbott, J. Abu-Threideh, R. Ahluwalia, E. Alexander, M. Alhawagri, J. Ali, J. Allen, M. Ames, S. Andrews, S. Angell, P. Antonacci, L. Antonacci-Fulton, B. Antoniou, J. Armstrong, C. Arnett, V. Atkins, K. Austin, C. Bailey, D. Baisden, B. Barbazuk, M. Barrett, L. Bartko, C. Bauer, H. Bauer, D. Baum, C. Beck, M. C. Becker, J. Bedell, K. Behymer, S. Behymer, E. Belter, G. Bemis, D. Bentley, A. Berghoff, K. Bernard, Z. Bevins, L. Bielicki, T. Biewald, L. Blackwood, R. Blaine, D. Blair, M. Blanchard, M. Blandford, D. Blasiar, J. Bolandis, S. Bolla, T. Bollinger, J. Bong, J. Boren-Prydydasz, S. Bourne, K. Bova, E. Boyer, K. Bradford, S. Brennan, M. Broy, D. Buatsi, C. Budnicki, M. Burkett, J. Burkhart, C. Buss, J. Butler, D. Caldwell, R. Caldwell, M. Cardenas, K. Carpenter, J. Carter, T. Carter, T. Carter, D. Casimere, A. Chapman, B. Chiapelli, A. T. Chinwalla, S. L. Chissoe, W. Christy, M. Cissell, B. Clark, M. J. Clark, K. Clarke, S. W. Clifton, J. Cloud, B. Coblitz, M. Cofman, M. Connell, J. Conyers, L. L. Cook, M. Cook, M. Cooper, V. Coppedge, M. Cordes, H. Cordum, M. Cotton, W. Courtney, K. Creason, J. M. Crockett, K. Crouse, T. Crum, M. Dante, R. Davenport, M. David, S. Davidson, T. Davidson, S. Davis, A. Delehaunty, K. D. Delehaunty, S. Dempsey, A. Desai, J. Despot, M. Dickes, K. Dickinson, N. Dietrich, G. Dignan, R. Dixon, A. Doebber, N. Doerr, M. Donoho, M. Dotson, J. Doucett, K. Drone, F. Du, H. Du, Z. Du, C. Dubbelde, G. Duckels, S. Eddy, S. Edinger, J. Edwards, T. Ehlman, J. Eldred, A. Elkin, G. Elliott, E. Exum, A. Falk, K. Farrow, A. Favello, J. Fedele, G. Fewell, D. Ficenec, T. Fiedler, L. Flagg, A. Fleming, N. Florence, J. Fries, W. Fronick, J. Fryman, D. Fuhrmann, L. A. Fulton, R. S. Fulton, D. Gaige, T. Gaige, J. Garrett, S. Gattung, C. Geisel, S. Geisel, A. Gibson, E. Gibson, C. Giddings, B. Gillam, Y. Gincherman, W. R. Gish, E. Glaser, D. Glossip, J. Godfrey, D. Goela, N. Goins, J. Gotway, E. Goyea-Gbadebo, L. Granderson, T. Graves, S. Gregory, S. Grewal, J. Griffin, H. Grover, G. Gualberto, C. Gund, W. Haakenson, K. Haglund, P. Hale, S. Hale, T. Hall, Z. Hamdan, C. Hannah, R. Harkins, G. Harmon, M. Harper, A. Harris, M. Harrison, R. Hart, K. Haub, J. Hawkins, C. Hawryszko, C. Heidbrink, K. Hendrix, J. Henkhaus, K. Henley, C. Henry, N. Hersberger, J. Heyen, M. Hickenbotham, P. Hill, T. Hillen, L. D. Hillier, K. Hinds, J. Hodges, E. Hoefgen, L. Holbrook, H. Hollingsworth, P. Holloway, M. Holman, A. Holmes, M. Hotic, S. Hou, S. Houshmandi, C. Howell, D. Hoyt, C. Hubbard, L. Isaiah, A. Isak, A. Jacobs, S. Jaeger, C. Jeliti, E. Jentes, A. Johnson, D. L. Johnson, B. Jones, K. Jones, R. Jones, C. Joshu, K. Kang, P. Kassos, K. Keen, J. Kellen, S. Kennedy, N. Keppler, M. Ketterman, K. Kim, S. Kitchell, D. Klebe, B. Klinke, J. Kloss, L. Knight, M. Koch, J. Kock, S. Kohlberg, I. Korg, D. Kovcic, J. Kraemer, J. B. Kramer, P. Krasucki, P. Krasucki, R. Krauss, C. Kremitzki, S. Kruchowski, T. Kucaba, M. Lacy, T. Lakanen, E. Lamar, K. Lane, Y. Langston, J. P. Latreille, D. Layman, T. Le, T. T. Pham Le, T. T. Le, J. J. Ledwith, N. Lee, L. Lehnert, S. Lennox, S. Leonard, K. Lesley, L. Levin, A. Levy, S. Lewis, L. Li, T. Littlejohn, N. Long, P. Lowery, S. Luxen, T. Lynch, J. Maas, J. MacDonald, L. Maggi, M. Maher, P. Marchetto, E. R. Mardis, C. Markovic, C. Marquis-Homeyer, M. A. Marra, G. Marth, J. C. Martin, J. Martin, S. Martinka, R. Maupin, K. Maxeiner, R. McAdow, M. C. M, C. McCabe, Q. McCray, B. McDill, K. McDonald, R. McDonald, T. McDonald, D. McDonough, R. McGrane, S. McKinney, M. McLellan, R. McMahon, J. D. McPherson, Y. McQuerrey, K. Mead, B. Meininger, B. Merry, R. Meyer, C. Meyers, K. Miller, N. Miller, W. Miller, T. L. Miner, B. Minges, P. J. Minx, S. Mishra, D. Moeller, L. Mohd Nor, K. Moire, B. Moore, T. Moore, R. Morales, N. Mudd, G. Mullen, M. Mullen, E. Mulvaney, J. Murray, M. Myers, A. Nash, W. Nash, J. Nelson, C. Nguyen, N. Nhan, C. Nicol, L. Niemann, L. Nothaker, T. Nwagbo, B. Oberkfell, D. O'Brien, D. O'Brien, T. Odunfa-Jones, M. K. Okuka, M. O'Malley, S. Owens, P. Ozersky, S. Page, D. Panussis, K. Pape, C. Parker, A. Pauley, E. Paulson, J. Peak, C. Pearman, D. Peluso, K. H. Pepin, D. Peterson, J. Pettiford, B. Pfeiffer, A. Phillips, G. Pierce, C. Pikula, A. Podhrasky, C. Pohl, T. Ponce, S. Puro, C. Ralph, J.

Randall, J. Randolph, J. Reed, A. Reily, D. Reiniesch, L. Reitz, J. Reskusich, C. Rhine, L. Rice, M. Richards, J. Richey, J. Rieff, J. Riley, E. Ritchey, J. Robertson, K. Robinson, S. Rock, T. Rohling, C. Rose, E. Ryan, J. Ryan, J. Ryan, S. Ryno, L. Sammons, B. Sandberg, T. Sandbothe, N. Sander, L. Sapetti, S. Sasso, M. Schaller, C. Schaus, D. Scheer, P. Scheet, E. Scherger, L. Schneider, B. Schultz, K. Scott, S. Scott, D. Scronce, R. Seim, M. Sekhon, S. Shafer, N. Shah, S. Shadid, K. Shapiro, P. Shelby, K. Shih, M. Slaughter, J. Small, A. Smith, A. Smith, E. Smith, J. Smith, N. Smith, R. Smith, B. Smoker, J. Snider, L. Spalding, J. Spieth, P. Steele, L. Stellyes, N. Stiziel, T. Stoneking, C. Strong, J. Strong, C. Strowmatt, E. Stuebe, J. Stumpf, R. Suk, H. Sun, C. Stutterer, G. Swift, S. Talcherkar, P. Thipkhosithkun, J. Thompson, A. M. Tin-Wollam, C. Tomlinson, M. Tonn, L. Trani, E. Trevaskis, S. Tucci, B. Twyman, K. Underwood, M. Ureta, P. Valencia, A. Van Bruntg, C. Veath, J. Veizer, C. Wagner-McPherson, J. Waligorski, C. Walker, R. Walker, T. Wall, J. Wallis, P. Wamsley, R. H. Waterston, P. Wedgeworth, A. Weihe, M. C. Wendl, N. Wheeler, S. White, N. Whitworth, D. Williams, A. Williamson, R. K. Wilson, K. Winchester, M. Winkelmann, J. Woessner, P. Wohldmann, J. Wolff, C. Wollam, K. Woods, J. P. Woolley, R. Worthington, X. Wu, K. Wylie, T. Wylie, M. Yandell, S. P. Yang, R. Yeh, M. Yoakum, S. Zerazion, X. Zheng, H. J. Zhu, M. Zidanic, A. Abrajano, A. Aerts, D. Alcivare, M. Altherr, G. Amico-Keller, J. Andora, C. H. Andredesz, T. Andriese, L. Arcaina, T. Arcaina, R. Archuleta, A. Arellano, N. Armstrong, L. Ashworth, C. Attix, A. Avery, A. Avila, J. Avila, H. Badri, M. Bakis, M. Balch, M. Banda, K. Beall, D. Beaton, J. Bercovitz, A. Bergmann, T. Beugelsdijk, T. Bobo, J. Boehm, M. Bondoc, M. P. Bonner, E. Bowen, W. Brannon, E. Branscomb, A. Brower, N. Brown, R. Brown, D. Bruce, R. Bruce, E. Brunkhorst, J. Bryant, J. Buckingham, K. Burkhart-Schultz, B. Bursell, M. Bussod, C. Campbell, E. Campbell, M. Campbell, C. Caoile, H. Cardenas, M. Cepeda, P. Chain, S. Chaparro, L. Chasteen, X. Chen, J. F. Cheng, S. G. Chin, C. Chinn, M. Christensen, A. Chung, R. Cifelli, L. Clark, J. Cofield, J. R. Cohn, R. Colayco, A. Copeland, R. Cordray, E. Connell, L. Corsetti, T. Critchlow, P. Critz, L. Danganan, W. Dean, L. Deaven, K. Deere, P. Dehal, Z. Deng, J. C. Detter, S. Detter, J. M. Dias, V. Dias, M. Dickson, R. DiGennaro, K. Dilts, M. Dimitrijevic-Bussod, K. Dixon, L. Do, N. Doggett, S. Duarte, C. Elkin, A. M. Erler, J. Fawcitt, J. Fey, M. Fink, K. Fischer, L. Fischer, J. P. Fitch, D. Flowers, P. Folta, D. Fotopulos, M. Fourcade, K. Frankel, M. Frazier, J. Fridlyand, S. Gammon, A. Georgescu, A. Geotina, I. Gil, T. Glavina, K. Golvineaux, S. Goodman, L. Goodwin, L. Gordon, K. Gould, B. Gray, L. Green, J. Griffith, J. Grimwood, M. Groza, H. Guarin, K. Gunning, C. Ha, C. Halsey, S. Hammond, C. Han, T. Hawkins, N. Henderson, W. Hom, R. Hosseini, Z. Huang, H. Hughes-Hull, D. Humphries, D. Hupman, J. Hurshman, K. Hutchings, D. Hyatt, J. Jaklevic, K. Jamaca, T. Jenecki, J. Jett, P. Jewitt, L. Jiang, J. Jin, M. Jones, E. Jung, K. Kadner, H. Kapur, L. Kegg, S. S. Khine, J. Kim, H. Kimball-Rojeski, W. Kimmerly, C. Ko, A. Kobayashi, W. Kolbe, K. Kommander, M. Krawczyk, B. Kronmiller, V. A. Krysiak, C. Kuhn, J. Lamerdin, M. Land, F. Larimer, B. Lato, J. H. Lee, M. Lee, K. Lehmann, T. Leyba, K. Lindo, K. Lindquist, A. Linkowski, K. Litton, S. Y. Liu, C. Llewellyn-Silva, R. Lobb, J. Logan, J. Longmire, J. L. Lopez, Y. Lou, S. Lowry, X. Lu, S. Lucas, M. Machrus, M. Macht, R. Madabhushi, R. Mahnke, M. Maltbie, M. Mariano, L. M. Marieiro, C. Martin, J. Martin, M. Martinez, P. McReady, P. McGurn, K. McMurry, C. Medina, K. Meier, L. Meincke, J. Menke, J. Meyne, T. Miguel, C. Miller, T. Milligan, S. Miner, V. Montgomery, D. Moy, M. Mundt, C. Munk, R. Mural, R. Myers, M. Narla, D. Nelson, J. Neunkirch, A. Newman, H. Nguyen, L. Nguyen, Q. Nguyen, M. Nolan, P. Oddone, J. Olivas, A. Olsen, D. Ow, M. Parang, B. Parson-Quintana, B. Patel, S. Patel, Y. Peng, Z. Peng, K. Peterman, B. Petit, J. Pfeiffer, H. Phan, S. Pitluck, L. Pittson, I. Plajzer-Frick, M. Pollard, P. Poundstone, E. Prakash, P. Predki, J. Primus, L. Probst, E. Prusso, G. Quan, L. Ramirez, M. Ramirez, D. Randolph, I. Rapier, W. Regala, C. Reiter, X. Ren, P. Richardson, D. Ricke, D. Robinson, J. Rodriguez, G. Sakaldasis, C. Sanders, R. Sarmiento, E. Saunders, D. Schmoyer, J. Schmutz, D. Scott, D. Scott, M. Shah, J. Shang, M. Shin, J. Shreve, J. Simoni, J. Sims, L. Sindelar, E. Skowronski, T. Slezak, J. Smith, J. Snoddy, G. Stanley, S. Stilwagen, L. Stubbs, J.

Stultz, S. Subramanian, R. Sutherland, K. Tacey, T. Takenaka, T. Tatum, A. Terry, J. Tesmer, J. Thiel, P. Thomas, L. Thompson, S. Thompson, W. Thompson, G. Tong, D. Torney, M. Tran, M. Trankiem, S. Trong, M. Y. Tsai, H. Turner, J. Turner, J. Turtrice, E. Uberbacher, C. Un, Q. Ung, R. V. Luchen, M. Vargas, N. P. Vartanian, N. P. Velasco, O. Velasquez, C. Vertuca, V. S. Viswanathan, J. Wagner, M. Wagner, W. Wan, M. Wang, E. Wehri, R. Weidenbach, S. Wenning, S. Wentz, C. White, J. White, S. White, A. Williams, D. Wilson, B. Winleblech-Kelly, J. R. Wollard, L. Woo, J. Woolley, T. Wright, M. Wycoff-Montegro, J. Yang, M. Yeh, C. Yu, B. Yumae, D. W. Zimmerman, C. Adams, B. Adio-Odulo, C. Allen, H. Allen, H. Amaratunge, A. Arenson, M. Bailey, T. Banks, J. Barbaria, K. Bimage, B. Bodota, D. Bonnin, J. Bouck, A. Brooks, E. Brown, M. J. Brown, N. Bryant, C. Buhay, P. Burch, C. Buckett, K. Burrell, T. Carron, K. Carter, S. Cavazos, J. Chacko, D. Chavez, G. Chen, M. Chen, R. Chen, Z. Chen, C. Christopoulos, K. Clerc-Blankenburg, R. Cockrell, C. Cox, M. Coyle, S. Dathorne, R. David, M. L. Davila, C. Davis, L. Davy-Carroll, O. Delgado, A. Denn, D. DeShazo, Y. Ding, H. Dinh, K. Douthwaite, H. Draper, S. Dugan-Rocha, K. J. Durbin, C. Earnhardt, D. Edgar, C. Allen, C. Elhaj, M. Escotto, T. Falls, N. Flagg, J. Forcum-Tansey, P. Foster, J. P. Frantz, A. Gabisi, A. Garcia, D. K. Garcia, T. Garner, R. A. Gibbs, R. Gill, J. H. Gorrell, L. L. Gorrell, W. Guevara, P. Gunaratne, K. Hamilton, V. Hanak, K. Harris, P. Havlak, A. Hawes, J. Hernandez, O. Hernandez, A. Hodgson, M. Hogues, B. Hollins, F. Homsi, H. Hosak, X. Hou, J. Huber, J. Hume, L. R. Jackson, Y. Jia, B. Johnson, R. Johnson, A. Jolivet, M. Jones, S. Joudah, S. Kaminsky, J. Kelly, S. Kelly, E. Karlsson, U. Khan, K. LaQuisha, N. Kondejewski, C. Kovar, J. Kratovic, N. Kondejewski, R. Kucherlapati, B. Leal, L. K. Lewis, L. Lewis, N. Kondejewski, J. Li, O. Lichtarge, J. Liu, W. Liu, L. Q. Logan, O. Logan, H. Loulseged, R. Lozado, J. Lu, A. Lucier, R. Lucier, T. Ly, J. Ma, M. Maheshwari, P. Mapua, R. Martin, A. Martindale, C. Martinez, E. Martinez, E. Massey, S. Mawhiney, M. McLeod, M. Meador, G. Mei, I. Mercado, M. Metzker, G. Miner, T. Mitchell, W. Mo, K. Mohabbat, B. Montgomery, K. Montgomery, M. Morgan, S. Morris, M. Moser, D. Muzny, S. Nash, S. Naylor, D. Neal, A. Nelson, D. Nelson, N. Newton, A. Nguyen, B. V. Nguyen, N. Nguyen, N. Nguyen, E. Nickerson, S. Nwokenkwo, M. Oguh, G. Okwuonu, G. Oswal, R. Oviedo, A. Pace, B. Parish, S. Paxton, B. Payton, L. Perez, L. Peters, A. Pickens, N. Pieper, E. Primus, L. L. Pu, M. Quiles, J. Quiroz, D. Reiter, Y. Ren, C. M. Rives, A. Rojas, S. J. Ruiz, G. Savery, S. Scherer, G. Scott, H. Shen, M. Simon, I. Sisson, E. Sodergren, T. Sonaike, L. Savage, A. Sparks, H. Stanley, H. Stone, H. Sutton, A. Svatek, L. A. Svetz, P. Tabor, K. Tamerisa, C. Taylor, T. Taylor, N. Thomas, S. Thomas, K. Timms, L. T. Tran, K. Usmani, L. Vasquez, V. Vera, D. Villalon, D. Villasana, Q. Vo, D. Walker, D. Walker, R. Wall, J. Wang, S. Wang, S. Ward-Moore, R. Warren, S. Watlington, M. Watrous, G. Weinstock, D. Wheeler, G. Williams, A. Williamson, R. Wleczyk, S. Wooden, K. Worley, G. Wrensford, J. Zhou, X. Zhou, S. Zorilla, T. Aizu, R. Arai, Y. Asahi, F. Ejima, A. Fujioka, A. Fujiyama, K. Fukano, R. Fukawa, Q. Gu, M. Hattori, M. Hirose, M. Horishima, K. Ishii, H. Ishizaki, E. Isozaki, N. Ito, T. Itoh, C. Kawagoe, Y. Kobayshi, N. Kodaka, M. Kondo, Y. Matsumura, Y. Mitani, H. Morita, A. Motoyama, S. Nagao, S. Nakagawa, K. Nakamura, C. Nakano, A. Nashida, Y. Odama, N. Omori, Y. Ono, K. Oshima, Y. Oyama, R. Ozawa, H. S. Park, R. Sakai, Y. Sakaki, H. Seki, H. Shimizu, J. Sun, T. Tahara, T. Tagaki, S. Takiguchi, M. Tanaka, R. Tanaka, T. Taylor, Y. Terada, M. Tochigi, N. Tomioka, Y. Totoki, A. Toyoda, Y. Tsukamoto, S. Tsukuni, R. Tsuzuki, N. Uyama, H. Wada, H. Watanabe, T. Yada, K. Yakushiji, N. Yamato, Y. Yamashita, S. Yokoyama, M. Yonezawa, S. Yoshida, F. Artiguenave, N. Barbe, M. Besnard, M. Besnard, D. Boscus, S. Briez, P. Brottier, T. Bruls, L. Cattolico, N. Cha, C. Da Silva, I. Dubois, M. Gouyvenoux, G. Gyapay, R. Heilig, S. Leclerc, M. Levy, G. Magdelenat, E. Pelletier, J. L. Petit, C. Robert, W. Saurin, B. Vacherie, V. Vico, J. Weissenbach, P. Wincker, M. Bakis, R. Bashirzadeh, J. Battles, M. Bodnaruk, G. Breton, J. Brown, C. Butler, P. Cahill, A. Caron, P. Daggett, T. Dorman, L. Doucette-Stamm, J. A. Dubois, N. Edwards, J. Ezedi, S. Flynn, L. Freeman, R. Gibson, D. Gleeson, G. Gryan, B. Herman, J. Hitti, T. Ho, K. Holtham, K. Huynh, C. Hynds, M. Johnson, P. Joseph, R. Kadel-Garcia, V. Kamath, A. Kana, K. Keane,

K. Kopcewiez, A. Lach, A. Lee, H. M. Lee, R. Little, W. Lumm, D. Madan, R. Magararu, J. I. Mao, L. Mitnik-Gankin, M. Munoz, M. Nguyen, W. Nielson, S. Prabhakar, J. Prescott-Roy, D. Qiu, B. Reinemann, S. Robinson, M. Roche, D. Rosetti, M. Rubenfield, O. Russakovskaya, J. Segal, D. R. Smith, P. Snell, M. Stroika, G. A. Turenne, J. Walsh, Y. Wang, K. Weinstock, G. Wheaton, L. Wong, Q. Xu, H. Yang, E. Zafiropoulos, E. Zhang, C. Baumgart, I. Baumgart, K. Blechschmidt, E. Boehm, C. Brunnckow, N. Creutzburg, M. Dette, B. Drescher, B. Drescher, P. Eibmann, S. Fabisch, B. Fischer, S. Foerste, P. Galgoczy, S. Gallert, G. Glockner, Y. Gorlich, C. Grosser, J. Hamann, I. Heintze, N. Jahn, E. Kantowski, H. Klabunde, S. Kluge, D. Lagemann, S. Landmann, R. Lehmann, D. Lenk, H. Ludewig, E. Meier, U. Menzel, E. Michaelis, K. Mockel, K. Mortag, O. Muller, G. Nordsiek, G. Nyakatura, B. Pawelka, U. Petz, U. Pick, M. Platzer, C. Pohlmann, A. Polley, B. Raguschke, N. Rahnis, K. Reichwald, A. Rosenthal, S. Rosenthal, S. Rothe, A. Rump, R. Schattevoy, A. Schauer, M. Schilhabel, M. Schilling, L. Schlenkert, M. L. Schmid, J. Schoemburg, A. Schudy, R. Schulz, S. Taudien, B. Tautkus, M. Teuchtler, B. Voigt, J. Weber, C. Wenderoth, D. Werler, T. Wiehe, N. Zeise, R. Zenker, W. D. Zimmerman, J. Bao, Q. Bao, W. Bao, S. Bi, X. Bian, L. Bolund, T. Cai, T. Cao, Y. Cao, B. Chen, C. Chen, J. Chen, J. Chen, J. Chen, T. Chen, Y. Chen, Z. Chen, H. Cui, J. Cui, P. Cui, L. Dai, H. Ding, H. Dong, W. Dong, X. Dong, Y. Du, H. Fan, J. Fang, H. Feng, J. Feng, X. Feng, G. Fu, J. Gao, Q. Gao, Y. Gao, J. Geng, G. Gong, J. Gong, J. Gu, W. Gu, X. Gu, Q. Guan, Q. Gui, D. Guo, F. He, J. He, L. He, J. Hu, S. Hu, F. Huang, G. Huang, J. Jia, N. Jia, L. Jiang, Y. Jin, Y. Jin, N. Kang, M. C. King, Y. Kong, M. Lei, C. Li, C. Li, E. Li, G. Li, J. Li, J. Li, L. Li, L. Li, M. Li, N. Li, R. Li, S. Li, S. Li, S. Li, S. Li, T. Li, W. Li, W. Li, Y. Li, Y. Li, Z. Li, J. Liao, W. Lin, B. Liu, H. Liu, K. Liu, N. Liu, S. Liu, W. Liu, X. Liu, Y. Liu, Y. Liu, Y. Liu, Z. Liu, T. Lu, Y. Lu, G. Lv, C. Ma, J. Ma, Q. Ma, S. Meng, F. Mu, Y. Niu, J. Pan, Q. Qi, X. Qi, X. Qian, Z. Qian, B. Qiang, Z. Qiao, S. Ren, L. Rong, Y. Shao, F. Shen, Y. Shen, Y. Shi, H. Shi, M. Smith, L. Song, S. Song, J. Sun, M. Sun, T. Sun, Y. Sun, Y. Sun, Y. Sun, W. Tan, X. Tan, X. Tang, R. Tao, Y. Tian, Y. Tian, J. Tong, Y. Tu, M. Wan, D. Wang, D. Wang, F. Wang, G. Wang, G. Wang, G. Wang, H. Wang, H. Wang, J. Wang, J. Wang, L. Wang, L. Wang, L. Wang, L. Wang, W. Wenjun, W. Xiaolei, X. Wang, X. Wang, Y. Wang, Y. Wang, C. Wu, D. Wu, Q. Wu, X. Wu, Y. Xi, F. Xie, R. Xu, S. Xu, W. Xu, Y. Xu, R. Xue, Y. Xue, C. Yan, F. Yan, G. Yan, Y. Huanming, S. Yang, X. Yang, Y. Bing, B. Yu, J. Yu, K. Yuan, Y. Zeng, D. Zhai, B. Zhang, F. Zhang, G. Zhang, G. Zhang, H. Zhang, L. Zhang, L. Zhang, L. Zhang, M. Zhang, M. Zhang, M. Zhang, R. Zhang, W. Zhang, X. Zhang, X. Zang, X. Zhang, Y. Zhang, Y. Zhang, Y. Zhang, Y. Zhang, Y. Zhang, H. Zhao, L. Zhao, Z. Zhao, Z. Zhen, M. Zhong, H. Zhou, N. Zhou, X. Zhou, Y. Zhou, Y. Zhou, B. Zhu, B. Zhu, G. Zhu, N. Zhu, Y. Zhu, Z. Zhu, N. Abassi, M. E. Ahearn, L. Baradarani, D. Baskin, B. Birditt, S. Bloom, C. Boysen, R. Bumgarner, R. Dickhoff, M. Dors, P. Fleetwood, C. Friedman, G. Harrison, L. Hood, R. James, A. Kaur, S. Lasky, I. Lee, C. Loretz, A. Madan, A. Madan, G. G. Mahairas, R. Nesbitt, S. Qin, A. Ratcliffe, L. Rowen, J. Seto, T. Shaffer, A. Smit, T. Smith, S. Swartzell, B. J. Trask, K. Wang, A. Federspiel, P. Foreman, S. Glukhov, N. Hansen, Z. Herman, R. Hyman, S. Kalman, O. Kurdi, J. Mao, R. Marathe, M. J. Proctor, A. Morehouse, P. Oefner, C. Palm, D. Ramirez, M. Rexan, M. Dela Rosa, M. Smith, D. Vollrath, J. Wilhelmy, T. Willils, S. Yu, E. Bajorek, C. Caoile, J. Carriere, D. R. Cox, M. Dickson, K. Dixon, L. Fischer, D. Flowers, D. Fotopulos, C. Garcia, D. Gold, J. Grimwood, L. Haydu, C. Holtzer, K. Litton, J. Logan, J. Lopez, C. Medina, R. M. Myers, L. Nguyen, L. Ramirez, A. Rodriquez, S. Rogers, A. Salazar, J. Schmutz, J. Shang, N. Stone, M. Tsai, O. Valesquez, S. Vartanian, D. Vitale, J. Wheeler, J. Yang, K. Bubb, R. Daza, C. Desmarais, S. Duenwald, K. Erickson, T. Gilbert, M. Hite, R. Hubley, W. Huges, S. Iodanoto, D. Jewett, C. Junker, A. Kas, R. Kaul, M. Le, R. Lim, L. Lytle, C. Magness, Z. Magness, M. Maza, E. McClelland, M. Olson, D. Passey, X. Q. Pham, K. A. Phelps, R. Qiu, S. Ramsey, C. Raymond, B. Richards, Z. Sadhegi, C. Saenphimmachak, E. Sims, A. Smit, M. Stone, T. Thomas, G. K. Wang, Z. Wu, J. Yu, Y. Zhou, N. Aoki, M. Asahina, S. Asakwa, K. Kawasaki, J. Kudoh, S. Minoshima, S. Mitsuyama, T. Sasaki, K. Shibuya, A. Shimizu, N. Shimizu, A. Shintani, Y. Yoshizaki, P. Aguayo, S. Arenare, D. Armstrong, M.

Athanasiou, M. Basit, D. Black, J. Brandon, J. Buettner, C. Butler, P. Card, S. Chamblis, J. Dunn, C. English, S. Ethridge, G. A. Evans, N. Federova, A. Fribish, M. Garza, M. Gordon, C. Gorman, O. D. Grant, L. Hahner, S. Hayes, J. Joslin, S. Lam, T. Le, T. Lester, E. Lewis, K. N. Loo, M. Loo, T. Major, J. McFairland, M. Nguyen, S. Osborne-Lawrence, I. Rakoshchik, J. Schageman, R. Schultz, S. Stimson, M. Tran, F. Varghese, N. Wagner, K. Waller, T. Ward, J. Wharton, J. Whitaker, J. N. Willcot, J. Zanoni, M. Ahmad, A. Bodenteich, F. Chen, L. Chu, J. Crabtree, S. Deschamps, A. Do, T. Do, J. Dolance, A. Dorman, C. Ducummon, A. Duty, M. Elraham, W. Elkins, F. Fang, Y. Fu, G. Hall, K. Hartman, K. Hill, P. Hu, X. Hu, A. Hua, E. Huang, J. Honggui, X. Jiang, S. Kenton, A. Khan, D. Kupfer, H. Lai, L. Lane, H. I. Lao, C. Lau, J. Lewis, S. Lewis, H. Li, S. Lin, P. Loh, E. Malaj, J. Milam, R. Morales-Diaz, F. Najar, T. Nguyen, Y. Ni, Oommen, H. Pan, B. Perry, S. Phan, S. Qi, Y. Qian, L. Ray, Q. Ren, Q. Ren, B. A. Roe, S. Shaull, D. Sloan, L. Song, J. Stone, J. Tian, R. Tian, Y. Tilahun, Q. Wang, Y. P. Wang, Z. Wang, D. White, J. White, D. Willingham, S. Wong, H. Wright, H. M. Wu, H. Wu, L. Yang, Z. Yao, Y. Yoon, M. Zhan, G. Zhang, L. Zhou, H. Zhu, S. Arndt, A. Beck, K. Borzym, D. Buczek, J. Chelly, F. Francis, K. Heitmann, S. Hennig, C. Hoff, E. Junker, P. Kioschis, S. Klages, M. Klein, A. Kosiura, M. Kube, I. Langer, H. Lehrach, S. Lehrack, I. Marquard, N. McDonell, A. Meindl, K. Moll, A. Monaco, A. Nemeth, A. Poustka, J. Ramser, R. Reinhardt, S. Schuelzchen, P. Seranski, P. Starke, C. Steffens, R. Sudbrak, K. Todd, M. L. Yaspo, M. de la Bastide, N. Dedhia, L. Gnoj, T. Gottesman, S. Granat, K. Haberman, A. Hameed, A. Hasegawa, J. Hoffman, E. Huang, K. Jenson, A. Johnson, N. Kaplan, M. Lodhi, A. Matero, W. R. McCombie, A. O'Shaughnessy, L. Parnell, R. Preston, M. Rodgiguez, K. Schutz, L. H. See, R. Shah, M. Shekher, N. Shohdy, L. Spiegel, I. Swaby, S. Till, D. Vil, H. Blocker, P. Brandt, A. Conrad, S. Dose, M. Grimm, K. Hornischer, D. Jarke, G. Kauer, H. Lohnert, G. Nordsiek, J. Reichelt, M. Scharfe, O. Schon, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, B. Birren, E. Birney, P. Bork, J. B. Bouck, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, A. T. Chinwalla, D. Church, M. Clamp, F. S. Collins, R. R. Copley, T. Doerks, R. Durbin, S. R. Eddy, E. E. Eichler, W. FitzHugh, A. Felsenfeld, T. S. Furey, J. Galagan, R. A. Gibbs, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermiakob, L. D. Hillier, K. Hokamp, T. Hubbard, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, E. S. Lander, T. M. Lowe, A. McLysaght, J. Mesirov, T. Mikkelsen, J. V. Moran, N. Mulder, J. C. Mullikn, C. Nusbaum, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Sulston, J. Szustakowki, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Waterson, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, M. C. Zody, R. Agarwala, L. Aravind, H. C. Chen, D. Church, W. Jang, P. Kitts, E. V. Koonin, G. Schuler, D. Thierry-Mieg, M. J. Thierry, L. Wagner, Y. I. Wolf, E. Birney, N. Reashci, P. Sterk, K. Fukami-Kobayashi, T. Gojobori, K. Ikeo, T. Imanishi, S. Myazaki, K. Nishikawa, M. Ohta, H. Sugawara, Y. Tateno, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. Wetterstrand, A. Patrinos and M. J. Morgan (2001). "Initial sequencing and analysis of the human genome." Nature 409(6822): 860-921.

Irion, S., H. Luche, P. Gadue, H. J. Fehling, M. Kennedy and G. Keller (2007). "Identification and targeting of the ROSA26 locus in human embryonic stem cells." Nat Biotechnol 25(12): 1477-1482.

Jacome, A., S. Navarro, J. A. Casado, P. Rio, L. Madero, J. Estella, J. Sevilla, I. Badell, J. J. Ortega, T. Olive, H. Hanenberg, J. C. Segovia and J. A. Bueren (2006). "A simplified approach to improve the efficiency and safety of ex vivo hematopoietic gene therapy in fanconi anemia patients." Hum Gene Ther 17(2): 245-250.

Ji, X., J. Kong and S. A. Liebhaber (2011). "An RNA-protein complex links enhanced nuclear 3' processing with cytoplasmic mRNA stabilization." EMBO J 30(13): 2622-2633.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Science 337(6096): 816-821.

Joenje, H., F. Arwert, A. W. Eriksson, H. de Koning and A. B. Oostra (1981). "Oxygen-dependence of chromosomal aberrations in Fanconi's anaemia." Nature 290(5802): 142-143.

Kamel-Reid, S. and J. E. Dick (1988). "Engraftment of immune-deficient mice with human hematopoietic stem cells." Science 242(4886): 1706-1709.

Kang, E. M., U. Choi, N. Theobald, G. Linton, D. A. Long Priel, D. Kuhns and H. L. Malech (2010). "Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils." Blood 115(4): 783-791.

Kang, H. J., C. C. Bartholomae, A. Paruzynski, A. Arens, S. Kim, S. S. Yu, Y. Hong, C. W. Joo, N. K. Yoon, J. W. Rhim, J. G. Kim, C. Von Kalle, M. Schmidt, S. Kim and H. S. Ahn (2011). "Retroviral gene therapy for X-linked chronic granulomatous disease: results from phase I/II trial." Mol Ther 19(11): 2092-2101.

Karginov, F. V. and G. J. Hannon (2010). "The CRISPR system: small RNA-guided defense in bacteria and archaea." Mol Cell 37(1): 7-19.

Kaufmann, K. B., H. Buning, A. Galy, A. Schambach and M. Grez (2013). "Gene therapy on the move." EMBO Mol Med 5(11): 1642-1661.

Kee, Y. and A. D. D'Andrea (2010). "Expanded roles of the Fanconi anemia pathway in preserving genomic stability." Genes Dev 24(16): 1680-1694.

Kee, Y. and A. D. D'Andrea (2012). "Molecular pathogenesis and clinical management of Fanconi anemia." J Clin Invest 122(11): 3799-3806.

Kelly, P. F., S. Radtke, C. von Kalle, B. Balcik, K. Bohn, R. Mueller, T. Schuesler, M. Haren, L. Reeves, J. A. Cancelas, T. Leemhuis, R. Harris, A. D. Auerbach, F. O. Smith, S. M. Davies and D. A. Williams (2007). "Stem cell collection and gene transfer in Fanconi anemia." Mol Ther 15(1): 211-219.

Kennedy, R. D. and A. D. D'Andrea (2005). "The Fanconi Anemia/BRCA pathway: new faces in the crowd." Genes Dev 19(24): 2925-2940.

Kim, H. and J. S. Kim (2014). "A guide to genome engineering with programmable nucleases." Nat Rev Genet 15(5): 321-334.

Kim, J. S., H. J. Lee and D. Carroll (2010). "Genome editing with modularly assembled zinc-finger nucleases." Nat Methods 7(2): 91; author reply 91-92.

Kim, Y., J. Kweon, A. Kim, J. K. Chon, J. Y. Yoo, H. J. Kim, S. Kim, C. Lee, E. Jeong, E. Chung, D. Kim, M. S. Lee, E. M. Go, H. J. Song, H. Kim, N. Cho, D. Bang, S. Kim and J. S. Kim (2013). "A library of TAL effector nucleases spanning the human genome." Nat Biotechnol 31(3): 251-258.

Kim, Y., J. Kweon and J. S. Kim (2013). "TALENs and ZFNs are associated with different mutation signatures." Nat Methods 10(3): 185.

Kim, Y., F. P. Lach, R. Desetty, H. Hanenberg, A. D. Auerbach and A. Smogorzewska (2011). "Mutations of the SLX4 gene in Fanconi anemia." Nat Genet 43(2): 142-146.

Kim, Y. G., J. Cha and S. Chandrasegaran (1996). "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain." Proc Natl Acad Sci U S A 93(3): 1156-1160.

Konermann, S., M. D. Brigham, A. E. Trevino, J. Joung, O. O. Abudayyeh, C. Barcena, P. D. Hsu, N. Habib, J. S. Gootenberg, H. Nishimasu, O. Nureki and F. Zhang (2015). "Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex." Nature 517(7536): 583-588.

Kotin, R. M., R. M. Linden and K. I. Berns (1992). "Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination." EMBO J 11(13): 5071-5078.

Kottemann, M. C. and A. Smogorzewska (2013). "Fanconi anaemia and the repair of Watson and Crick DNA crosslinks." Nature 493(7432): 356-363.

LaFountaine, J. S., K. Fathe and H. D. Smyth (2015). "Delivery and therapeutic applications of gene editing technologies ZFNs, TALENs, and CRISPR/Cas9." Int J Pharm 494(1): 180-194.

Larochelle, A., J. Vormoor and H. Heanenberg (1996). "Identification of primitive human hematopoietic cells capable of repopulating NOD7SCID mouse bone marrow: implications for gene therapy." Nature Medicine 2: 1329-1337.

Levitus, M., Q. Waisfisz, B. C. Godthelp, Y. de Vries, S. Hussain, W. W. Wiegant, E. Elghalbzouri-Maghrani, J. Steltenpool, M. A. Rooimans, G. Pals, F. Arwert, C. G. Mathew, M. Z. Zdzienicka, K. Hiom, J. P. De Winter and H. Joenje (2005). "The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J." Nat Genet 37(9): 934-935.

Li, H. L., N. Fujimoto, N. Sasakawa, S. Shirai, T. Ohkame, T. Sakuma, M. Tanaka, N. Amano, A. Watanabe, H. Sakurai, T. Yamamoto, S. Yamanaka and A. Hotta (2015). "Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9." Stem Cell Reports 4(1): 143-154.

Li, T., S. Huang, X. Zhao, D. A. Wright, S. Carpenter, M. H. Spalding, D. P. Weeks and B. Yang (2011). "Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes." Nucleic Acids Res 39(14): 6315-6325.

Lim, E. T., P. Wurtz, A. S. Havulinna, P. Palta, T. Tukiainen, K. Rehnstrom, T. Esko, R. Magi, M. Inouye, T. Lappalainen, Y. Chan, R. M. Salem, M. Lek, J. Flannick, X. Sim, A. Manning, C. Ladenvall, S. Bumpstead, E. Hamalainen, K. Aalto, M. Maksimow, M. Salmi, S. Blankenberg, D. Ardissino, S. Shah, B. Horne, R. McPherson, G. K. Hovingh, M. P. Reilly, H. Watkins, A. Goel, M. Farrall, D. Girelli, A. P. Reiner, N. O. Stitziel, S. Kathiresan, S. Gabriel, J. C. Barrett, T. Lehtimaki, M. Laakso, L. Groop, J. Kaprio, M. Perola, M. I. McCarthy, M. Boehnke, D. M. Altshuler, C. M. Lindgren, J. N. Hirschhorn, A. Metspalu, N. B. Freimer, T. Zeller, S. Jalkanen, S. Koskinen, O. Raitakari, R. Durbin, D. G. MacArthur, V. Salomaa, S. Ripatti, M. J. Daly, A. Palotie and P. Sequencing Initiative Suomi (2014). "Distribution and medical impact of loss-of-function variants in the Finnish founder population." PLoS Genet 10(7): e1004494.

Lim, J. K., W. G. Glass, D. H. McDermott and P. M. Murphy (2006). "CCR5: no longer a "good for nothing" gene--chemokine control of West Nile virus infection." Trends Immunol 27(7): 308-312.

Liu, G. H., K. Suzuki, M. Li, J. Qu, N. Montserrat, C. Tarantino, Y. Gu, F. Yi, X. Xu, W. Zhang, S. Ruiz, N. Plongthongkum, K. Zhang, S. Masuda, E. Nivet, Y. Tsunekawa, R. D. Soligalla, A. Goebl, E. Aizawa, N. Y. Kim, J. Kim, I. Dubova, Y. Li, R. Ren, C. Benner, A. del Sol, J. Bueren, J. P. Trujillo, J. Surralles, E. Cappelli, C. Dufour, C. R. Esteban and J. C. Izpisua Belmonte (2014). "Modelling Fanconi anemia pathogenesis and therapeutics using integration-free patient-derived iPSCs." Nat Commun 5: 4330.

Liu, J. M., S. Kim, E. J. Read, M. Futaki, I. Dokal, C. S. Carter, S. F. Leitman, M. Pensiero, N. S. Young and C. E. Walsh (1999). "Engraftment of hematopoietic progenitor cells transduced with the Fanconi anemia group C gene (FANCC)." Hum Gene Ther 10(14): 2337-2346.

Liu, R., W. A. Paxton, S. Choe, D. Ceradini, S. R. Martin, R. Horuk, M. E. MacDonald, H. Stuhlmann, R. A. Koup and N. R. Landau (1996). "Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection." Cell 86(3): 367-377.

Lo Ten Foe, J. R., M. A. Rooimans, L. Bosnoyan-Collins, N. Alon, M. Wijker, L. Parker, J. Lightfoot, M. Carreau, D. F. Callen, A. Savoia, N. C. Cheng, C. G. van Berkel, M. H. Strunk, J. J. Gille, G. Pals, F. A. Kruyt, J. C. Pronk, F. Arwert, M. Buchwald and H. Joenje (1996). "Expression cloning of a cDNA for the major Fanconi anaemia gene, FAA." Nat Genet 14(3): 320-323.

Lombardo, A., D. Cesana, P. Genovese, B. Di Stefano, E. Provasi, D. F. Colombo, M. Neri, Z. Magnani, A. Cantore, P. Lo Riso, M. Damo, O. M. Pello, M. C. Holmes, P. D. Gregory, A. Gritti, V. Broccoli, C. Bonini and L. Naldini (2011). "Site-specific integration and tailoring of cassette design for sustainable gene transfer." Nat Methods 8(10): 861-869.

Lombardo, A., P. Genovese, C. M. Beausejour, S. Colleoni, Y. L. Lee, K. A. Kim, D. Ando, F. D. Urnov, C. Galli, P. D. Gregory, M. C. Holmes and L. Naldini (2007). "Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery." Nat Biotechnol 25(11): 1298-1306.

Ma, N., B. Liao, H. Zhang, L. Wang, Y. Shan, Y. Xue, K. Huang, S. Chen, X. Zhou, Y. Chen, D. Pei and G. Pan (2013). "Transcription activator-like effector nuclease (TALEN)-mediated gene correction in integration-free beta-thalassemia induced pluripotent stem cells." J Biol Chem 288(48): 34671-34679.

MacMillan, M. L., T. E. DeFor, J. A. Young, K. E. Dusenbery, B. R. Blazar, A. Slungaard, H. Zierhut, D. J. Weisdorf and J. E. Wagner (2015). "Alternative donor hematopoietic cell transplantation for Fanconi anemia." Blood 125(24): 3798-3804.

MacMillan, M. L. and J. E. Wagner (2010). "Haematopoeitic cell transplantation for Fanconi anaemia - when and how?" Br J Haematol 149(1): 14-21.

Maeder, M. L., S. J. Linder, V. M. Cascio, Y. Fu, Q. H. Ho and J. K. Joung (2013). "CRISPR RNA-guided activation of endogenous human genes." Nat Methods 10(10): 977-979.

Maetzig, T., M. H. Brugman, S. Bartels, N. Heinz, O. S. Kustikova, U. Modlich, Z. Li, M. Galla, B. Schiedlmeier, A. Schambach and C. Baum (2011). "Polyclonal fluctuation of lentiviral vector-transduced and expanded murine hematopoietic stem cells." Blood 117(11): 3053-3064.

Maggio, I. and M. A. Goncalves (2015). "Genome editing at the crossroads of delivery, specificity, and fidelity." Trends Biotechnol 33(5): 280-291.

Mak, A. N., P. Bradley, R. A. Cernadas, A. J. Bogdanove and B. L. Stoddard (2012). "The crystal structure of TAL effector PthXo1 bound to its DNA target." Science 335(6069): 716-719.

Makarova, K. S., D. H. Haft, R. Barrangou, S. J. Brouns, E. Charpentier, P. Horvath, S. Moineau, F. J. Mojica, Y. I. Wolf, A. F. Yakunin, J. van der Oost and E. V. Koonin (2011). "Evolution and classification of the CRISPR-Cas systems." Nat Rev Microbiol 9(6): 467-477.

Malech, H. L., P. B. Maples, N. Whiting-Theobald, G. F. Linton, S. Sekhsaria, S. J. Vowells, F. Li, J. A. Miller, E. DeCarlo, S. M. Holland, S. F. Leitman, C. S. Carter, R. E. Butz, E. J. Read, T. A. Fleisher, R. D. Schneiderman, D. E. Van Epps, S. K. Spratt, C. A. Maack, J. A. Rokovich, L. K. Cohen and J. I. Gallin (1997). "Prolonged production of NADPH oxidase-corrected granulocytes after gene therapy of chronic granulomatous disease." Proc Natl Acad Sci U S A 94(22): 12133-12138.

Mali, P., J. Aach, P. B. Stranges, K. M. Esvelt, M. Moosburner, S. Kosuri, L. Yang and G. M. Church (2013). "CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering." Nat Biotechnol 31(9): 833-838.

Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville and G. M. Church (2013). "RNA-guided human genome engineering via Cas9." Science 339(6121): 823-826.

Masat, E., G. Pavani and F. Mingozzi (2013). "Humoral immunity to AAV vectors in gene therapy: challenges and potential solutions." Discov Med 15(85): 379-389.

McCune, J. M., R. Namikawa, H. Kaneshima, L. D. Shultz, M. Lieberman and I. L. Weissman (1988). "The SCID-hu mouse: murine model for the analysis of human hematolymphoid differentiation and function." Science 241(4873): 1632-1639.

Meetei, A. R., J. P. de Winter, A. L. Medhurst, M. Wallisch, Q. Waisfisz, H. J. van de Vrugt, A. B. Oostra, Z. Yan, C. Ling, C. E. Bishop, M. E. Hoatlin, H. Joenje and W. Wang (2003). "A novel ubiquitin ligase is deficient in Fanconi anemia." Nat Genet 35(2): 165-170.

Meetei, A. R., M. Levitus, Y. Xue, A. L. Medhurst, M. Zwaan, C. Ling, M. A. Rooimans, P. Bier, M. Hoatlin, G. Pals, J. P. de Winter, W. Wang and H. Joenje (2004). "X-linked inheritance of Fanconi anemia complementation group B." Nat Genet 36(11): 1219-1224.

Meetei, A. R., A. L. Medhurst, C. Ling, Y. Xue, T. R. Singh, P. Bier, J. Steltenpool, S. Stone, I. Dokal, C. G. Mathew, M. Hoatlin, H. Joenje, J. P. de Winter and W. Wang (2005). "A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M." Nat Genet 37(9): 958-963.

Modlich, U., J. Bohne, M. Schmidt, C. von Kalle, S. Knoss, A. Schambach and C. Baum (2006). "Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity." Blood 108(8): 2545-2553.

Modlich, U., S. Navarro, D. Zychlinski, T. Maetzig, S. Knoess, M. H. Brugman, A. Schambach, S. Charrier, A. Galy, A. J. Thrasher, J. Bueren and C. Baum (2009). "Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors." Mol Ther 17(11): 1919-1928.

Montini, E., D. Cesana, M. Schmidt, F. Sanvito, C. C. Bartholomae, M. Ranzani, F. Benedicenti, L. S. Sergi, A. Ambrosi, M. Ponzoni, C. Doglioni, C. Di Serio, C. von Kalle and L. Naldini (2009). "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 119(4): 964-975.

Montini, E., D. Cesana, M. Schmidt, F. Sanvito, M. Ponzoni, C. Bartholomae, L. Sergi Sergi, F. Benedicenti, A. Ambrosi, C. Di Serio, C. Doglioni, C. von Kalle and L. Naldini (2006). "Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration." Nat Biotechnol 24(6): 687-696.

Moscou, M. J. and A. J. Bogdanove (2009). "A simple cipher governs DNA recognition by TAL effectors." Science 326(5959): 1501.

Muller, L. U., M. D. Milsom, C. E. Harris, R. Vyas, K. M. Brumme, K. Parmar, L. A. Moreau, A. Schambach, I. H. Park, W. B. London, K. Strait, T. Schlaeger, A. L. Devine, E. Grassman, A. D'Andrea, G. Q. Daley and D. A. Williams (2012). "Overcoming reprogramming resistance of Fanconi anemia cells." Blood 119(23): 5449-5457.

Mussolino, C., T. Mlambo and T. Cathomen (2015). "Proven and novel strategies for efficient editing of the human genome." Curr Opin Pharmacol 24: 105-112.

Mussolino, C., R. Morbitzer, F. Lutge, N. Dannemann, T. Lahaye and T. Cathomen (2011). "A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity." Nucleic Acids Res 39(21): 9283-9293.

Nakanishi, K., F. Cavallo, L. Perrouault, C. Giovannangeli, M. E. Moynahan, M. Barchi, E. Brunet and M. Jasin (2011). "Homology-directed Fanconi anemia pathway cross-link repair is dependent on DNA replication." Nat Struct Mol Biol 18(4): 500-503.

Nakayama, T., M. B. Fish, M. Fisher, J. Oomen-Hajagos, G. H. Thomsen and R. M. Grainger (2013). "Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in Xenopus tropicalis." Genesis 51(12): 835-843.

Naldini, L. (2011). "Ex vivo gene transfer and correction for cell-based therapies." Nat Rev Genet 12(5): 301-315.

Navarrete, C. and M. Contreras (2009). "Cord blood banking: a historical perspective." Br J Haematol 147(2): 236-245.

Navarro, S., P. Rio and J. Bueren (2015). "Perspectives on gene therapy for Fanconi anemia."

Neveling, K., D. Endt, H. Hoehn and D. Schindler (2009). "Genotype-phenotype correlations in Fanconi anemia." Mutat Res 668(1-2): 73-91.

Nienhuis, A. W., C. E. Dunbar and B. P. Sorrentino (2006). "Genotoxicity of retroviral integration in hematopoietic cells." Mol Ther 13(6): 1031-1049.

Niu, Y., B. Shen, Y. Cui, Y. Chen, J. Wang, L. Wang, Y. Kang, X. Zhao, W. Si, W. Li, A. P. Xiang, J. Zhou, X. Guo, Y. Bi, C. Si, B. Hu, G. Dong, H. Wang, Z. Zhou, T. Li, T. Tan, X. Pu, F. Wang, S. Ji, Q. Zhou, X. Huang, W. Ji and J. Sha (2014). "Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos." Cell 156(4): 836-843.

Osborn, M. J., R. Gabriel, B. R. Webber, A. P. DeFeo, A. N. McElroy, J. Jarjour, C. G. Starker, J. E. Wagner, J. K. Joung, D. F. Voytas, C. von Kalle, M. Schmidt, B. R. Blazar and J. Tolar (2015). "Fanconi anemia gene editing by the CRISPR/Cas9 system." Hum Gene Ther 26(2): 114-126.

Ott, M. G., M. Schmidt, K. Schwarzwaelder, S. Stein, U. Siler, U. Koehl, H. Glimm, K. Kuhlcke, A. Schilz, H. Kunkel, S. Naundorf, A. Brinkmann, A. Deichmann, M. Fischer, C. Ball, I. Pilz, C. Dunbar, Y. Du, N. A. Jenkins, N. G. Copeland, U. Luthi, M. Hassan, A. J. Thrasher, D. Hoelzer, C. von Kalle, R. Seger and M. Grez (2006). "Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1." Nat Med 12(4): 401-409.

Pagano, G., A. A. Talamanca, G. Castello, M. d'Ischia, F. V. Pallardo, S. Petrovic, B. Porto, L. Tiano and A. Zatterale (2013). "From clinical description, to in vitro and animal studies, and backward to patients: oxidative stress and mitochondrial dysfunction in Fanconi anemia." Free Radic Biol Med 58: 118-125.

Pagano, G., A. A. Talamanca, G. Castello, F. V. Pallardo, A. Zatterale and P. Degan (2012). "Oxidative stress in Fanconi anaemia: from cells and molecules towards prospects in clinical management." Biol Chem 393(1-2): 11-21.

Pallardo, F. V., A. Lloret, M. Lebel, M. d'Ischia, V. C. Cogger, D. G. Le Couteur, M. N. Gadaleta, G. Castello and G. Pagano (2010). "Mitochondrial dysfunction in some oxidative stress-related genetic diseases: Ataxia-Telangiectasia, Down Syndrome, Fanconi Anaemia and Werner Syndrome." Biogerontology 11(4): 401-419.

Papapetrou, E. P., G. Lee, N. Malani, M. Setty, I. Riviere, L. M. Tirunagari, K. Kadota, S. L. Roth, P. Giardina, A. Viale, C. Leslie, F. D. Bushman, L. Studer and M. Sadelain (2010). "Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells." Nat Biotechnol 29(1): 73-78.

Park, C. Y., J. Kim, J. Kweon, J. S. Son, J. S. Lee, J. E. Yoo, S. R. Cho, J. H. Kim, J. S. Kim and D. W. Kim (2014). "Targeted inversion and reversion of the blood coagulation factor 8 gene in human iPS cells using TALENs." Proc Natl Acad Sci U S A 111(25): 9253-9258.

Pattanayak, V., C. L. Ramirez, J. K. Joung and D. R. Liu (2011). "Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection." Nat Methods 8(9): 765-770.

Peffault de Latour, R., R. Porcher, J. H. Dalle, M. Aljurf, E. T. Korthof, J. Svahn, R. Willemze, C. Barrenetxea, V. Mialou, J. Soulier, M. Ayas, R. Oneto, A. Bacigalupo, J. C. Marsh, C. Peters, G. Socie, C. Dufour, F. A. C. o. t. S. A. A. W. Party, B. Pediatric Working Party of the European Group for and T. Marrow (2013). "Allogeneic hematopoietic stem cell transplantation in Fanconi anemia: the European Group for Blood and Marrow Transplantation experience." Blood 122(26): 4279-4286.

Perez-Pinera, P., D. D. Kocak, C. M. Vockley, A. F. Adler, A. M. Kabadi, L. R. Polstein, P. I. Thakore, K. A. Glass, D. G. Ousterout, K. W. Leong, F. Guilak, G. E. Crawford, T. E. Reddy and C. A. Gersbach (2013). "RNA-guided gene activation by CRISPR-Cas9-based transcription factors." Nat Methods 10(10): 973-976.

Perez, E. E., J. Wang, J. C. Miller, Y. Jouvenot, K. A. Kim, O. Liu, N. Wang, G. Lee, V. V. Bartsevich, Y. L. Lee, D. Y. Guschin, I. Rupniewski, A. J. Waite, C. Carpenito, R. G. Carroll, J. S. Orange, F. D. Urnov, E. J. Rebar, D. Ando, P. D. Gregory, J. L. Riley, M. C. Holmes and C. H. June (2008). "Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases." Nat Biotechnol 26(7): 808-816.

Platt, R. J., S. Chen, Y. Zhou, M. J. Yim, L. Swiech, H. R. Kempton, J. E. Dahlman, O. Parnas, T. M. Eisenhaure, M. Jovanovic, D. B. Graham, S. Jhunjhunwala, M. Heidenreich, R. J. Xavier, R. Langer, D. G. Anderson, N. Hacohen, A. Regev, G. Feng, P. A. Sharp and F. Zhang (2014). "CRISPR-Cas9 knockin mice for genome editing and cancer modeling." Cell 159(2): 440-455.

Poirot, L., B. Philip, C. Schiffer-Mannioui, D. Le Clerre, I. Chion-Sotinel, S. Derniame, P. Potrel, C. Bas, L. Lemaire, R. Galetto, C. Lebuhotel, J. Eyquem, G. W. Cheung, A. Duclert, A. Gouble, S. Arnould, K. Peggs, M. Pule, A. M. Scharenberg and J. Smith (2015). "Multiplex Genome-Edited T-cell Manufacturing Platform for "Off-the-Shelf" Adoptive T-cell Immunotherapies." Cancer Res 75(18): 3853-3864.

Porteus, M. H. and D. Baltimore (2003). "Chimeric nucleases stimulate gene targeting in human cells." Science 300(5620): 763.

Porteus, M. H. and D. Carroll (2005). "Gene targeting using zinc finger nucleases." Nat Biotechnol 23(8): 967-973.

Pruett-Miller, S. M., D. W. Reading, S. N. Porter and M. H. Porteus (2009). "Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels." PLoS Genet 5(2): e1000376.

Pulliam, A. C., M. J. Hobson, S. L. Ciccone, Y. Li, S. Chen, E. F. Srour, F. C. Yang, H. E. Broxmeyer and D. W. Clapp (2008). "AMD3100 synergizes with G-CSF to mobilize repopulating stem cells in Fanconi anemia knockout mice." Exp Hematol 36(9): 1084-1090.

Ramirez, C. L., J. E. Foley, D. A. Wright, F. Muller-Lerch, S. H. Rahman, T. I. Cornu, R. J. Winfrey, J. D. Sander, F. Fu, J. A. Townsend, T. Cathomen, D. F. Voytas and J. K. Joung (2008). "Unexpected failure rates for modular assembly of engineered zinc fingers." Nat Methods 5(5): 374-375.

Ravera, S., D. Vaccaro, P. Cuccarolo, M. Columbaro, C. Capanni, M. Bartolucci, I. Panfoli, A. Morelli, C. Dufour, E. Cappelli and P. Degan (2013). "Mitochondrial respiratory chain Complex I defects in Fanconi anemia complementation group A." Biochimie 95(10): 1828-1837.

Raya, A., I. Rodriguez-Piza, G. Guenechea, R. Vassena, S. Navarro, M. J. Barrero, A. Consiglio, M. Castella, P. Rio, E. Sleep, F. Gonzalez, G. Tiscornia, E. Garreta, T. Aasen, A. Veiga, I. M. Verma, J. Surralles, J. Bueren and J. C. Izpisua Belmonte (2009). "Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells." Nature 460(7251): 53-59.

Reid, S., D. Schindler, H. Hanenberg, K. Barker, S. Hanks, R. Kalb, K. Neveling, P. Kelly, S. Seal, M. Freund, M. Wurm, S. D. Batish, F. P. Lach, S. Yetgin, H. Neitzel, H. Ariffin, M. Tischkowitz, C. G. Mathew, A. D. Auerbach and N. Rahman (2007). "Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer." Nat Genet 39(2): 162-164.

Rickman, K. A., F. P. Lach, A. Abhyankar, F. X. Donovan, E. M. Sanborn, J. A. Kennedy, C. Sougnez, S. B. Gabriel, O. Elemento, S. C. Chandrasekharappa, D. Schindler, A. D. Auerbach and A. Smogorzewska (2015). "Deficiency of UBE2T, the E2 Ubiquitin Ligase Necessary for FANCD2 and FANCI Ubiquitination, Causes FA-T Subtype of Fanconi Anemia." Cell Rep 12(1): 35-41.

Rio, P., R. Banos, A. Lombardo, O. Quintana-Bustamante, L. Alvarez, Z. Garate, P. Genovese, E. Almarza, A. Valeri, B. Diez, S. Navarro, Y. Torres, J. P. Trujillo, R. Murillas, J. C. Segovia, E. Samper, J. Surralles, P. D. Gregory, M. C. Holmes, L. Naldini and J. A. Bueren (2014). "Targeted gene therapy and cell reprogramming in Fanconi anemia." EMBO Mol Med 6(6): 835-848.

Rio, P., N. W. Meza, A. Gonzalez-Murillo, S. Navarro, L. Alvarez, J. Surralles, M. Castella, G. Guenechea, J. C. Segovia, H. Hanenberg and J. A. Bueren (2008). "In vivo proliferation advantage of genetically corrected hematopoietic stem cells in a mouse model of Fanconi anemia FA-D1." Blood 112(13): 4853-4861.

Rio, P., S. Navarro and J. Bueren From the molecular biology to the gene therapy of a DNA repair syndrome: Fanconi anemia, DNA repair and Human Health.

Rottman, J. B., K. P. Ganley, K. Williams, L. Wu, C. R. Mackay and D. J. Ringler (1997). "Cellular localization of the chemokine receptor CCR5. Correlation to cellular targets of HIV-1 infection." Am J Pathol 151(5): 1341-1351.

Rouet, P., F. Smih and M. Jasin (1994). "Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells." Proc Natl Acad Sci U S A 91(13): 6064-6068.

Sadelain, M., E. P. Papapetrou and D. Bushman (2012). "Safe harbors for the integration of new DNA in the human genome." Nature 12.

Saito, H., A. T. Hammond and R. E. Moses (1993). "Hypersensitivity to oxygen is a uniform and secondary defect in Fanconi anemia cells." Mutat Res 294(3): 255-262.

Sander, J. D., E. J. Dahlborg, M. J. Goodwin, L. Cade, F. Zhang, D. Cifuentes, S. J. Curtin, J. S. Blackburn, S. Thibodeau-Beganny, Y. Qi, C. J. Pierick, E. Hoffman, M. L. Maeder, C. Khayter, D. Reyon, D. Dobbs, D. M. Langenau, R. M. Stupar, A. J. Giraldez, D. F. Voytas, R. T. Peterson, J. R. Yeh and J. K. Joung (2011). "Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA)." Nat Methods 8(1): 67-69.

Sander, J. D. and J. K. Joung (2014). "CRISPR-Cas systems for editing, regulating and targeting genomes." Nat Biotechnol 32(4): 347-355.

Sawyer, S. L., L. Tian, M. Kahkonen, J. Schwartzentruber, M. Kircher, G. University of Washington Centre for Mendelian, F. C. Consortium, J. Majewski, D. A. Dyment, A. M. Innes, K. M. Boycott, L. A. Moreau, J. S. Moilanen and R. A. Greenberg (2015). "Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype." Cancer Discov 5(2): 135-142.

Scheckenbach, K., M. Morgan, J. Filger-Brillinger, M. Sandmann, B. Strimling, W. Scheurlen, D. Schindler, U. Gobel and H. Hanenberg (2012). "Treatment of the bone marrow failure in Fanconi anemia patients with danazol." Blood Cells Mol Dis 48(2): 128-131.

Schuler, D., A. Kiss and F. Fabian (1969). "Chromosomal peculiarities and "in vitro" examinations in Fanconi's anaemia." Humangenetik 7(4): 314-322.

Schwab, R. A., A. N. Blackford and W. Niedzwiedz (2010). "ATR activation and replication fork restart are defective in FANCM-deficient cells." EMBO J 29(4): 806-818.

Silva, G., L. Poirot, R. Galetto, J. Smith, G. Montoya, P. Duchateau and F. Paques (2011). "Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy." Curr Gene Ther 11(1): 11-27.

Silva, G. H., M. Belfort, W. Wende and A. Pingoud (2006). "From monomeric to homodimeric endonucleases and back: engineering novel specificity of LAGLIDADG enzymes." J Mol Biol 361(4): 744-754.

Singh, T. R., S. T. Bakker, S. Agarwal, M. Jansen, E. Grassman, B. C. Godthelp, A. M. Ali, C. H. Du, M. A. Rooimans, Q. Fan, K. Wahengbam, J. Steltenpool, P. R. Andreassen, D. A. Williams, H. Joenje, J. P. de Winter and A. R. Meetei (2009). "Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M." Blood 114(1): 174-180.

Singh, T. R., D. Saro, A. M. Ali, X. F. Zheng, C. H. Du, M. W. Killen, A. Sachpatzidis, K. Wahengbam, A. J. Pierce, Y. Xiong, P. Sung and A. R. Meetei (2010). "MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM." Mol Cell 37(6): 879-886.

Skinner, A. M., S. L. O'Neill, M. Grompe and P. Kurre (2008). "CXCR4 induction in hematopoietic progenitor cells from Fanca(-/-), -c(-/-), and -d2(-/-) mice." Exp Hematol 36(3): 273-282.

Skipper, K. A. and J. G. Mikkelsen (2015). "Delivering the Goods for Genome Engineering and Editing." Hum Gene Ther 26(8): 486-497.

Smih, F., P. Rouet, P. J. Romanienko and M. Jasin (1995). "Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells." Nucleic Acids Res 23(24): 5012-5019.

Smit, A. F. (1999). "Interspersed repeats and other mementos of transposable elements in mammalian genomes." Curr Opin Genet Dev 9(6): 657-663.

Smith, J. R., S. Maguire, L. A. Davis, M. Alexander, F. Yang, S. Chandran, C. ffrench-Constant and R. A. Pedersen (2008). "Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration." Stem Cells 26(2): 496-504.

Smogorzewska, A., S. Matsuoka, P. Vinciguerra, E. R. McDonald, 3rd, K. E. Hurov, J. Luo, B. A. Ballif, S. P. Gygi, K. Hofmann, A. D. D'Andrea and S. J. Elledge (2007). "Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair." Cell 129(2): 289-301.

Stein, S., M. G. Ott, S. Schultze-Strasser, A. Jauch, B. Burwinkel, A. Kinner, M. Schmidt, A. Kramer, J. Schwable, H. Glimm, U. Koehl, C. Preiss, C. Ball, H. Martin, G. Gohring, K. Schwarzwaelder, W. K. Hofmann, K. Karakaya, S. Tchatchou, R. Yang, P. Reinecke, K. Kuhlcke, B. Schlegelberger, A. J. Thrasher, D. Hoelzer, R. Seger, C. von Kalle and M. Grez (2010). "Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease." Nat Med 16(2): 198-204.

Stephan, V., V. Wahn, F. Le Deist, U. Dirksen, B. Broker, I. Muller-Fleckenstein, G. Horneff, H. Schroten, A. Fischer and G. de Saint Basile (1996). "Atypical X-linked severe combined immunodeficiency due to possible spontaneous reversion of the genetic defect in T cells." N Engl J Med 335(21): 1563-1567.

Strathdee, C. A., H. Gavish, W. R. Shannon and M. Buchwald (1992). "Cloning of cDNAs for Fanconi's anaemia by functional complementation." Nature 356(6372): 763-767.

Suzuki, N. M., A. Niwa, M. Yabe, A. Hira, C. Okada, N. Amano, A. Watanabe, K. Watanabe, T. Heike, M. Takata, T. Nakahata and M. K. Saito (2015). "Pluripotent cell models of fanconi anemia identify the early pathological defect in human hemoangiogenic progenitors." Stem Cells Transl Med 4(4): 333-338.

Svahn, J. and C. Dufour (2011). "Fanconi anemia - learning from children." Pediatr Rep 3 Suppl 2: e8.

Szymczak, A. L., C. J. Workman, Y. Wang, K. M. Vignali, S. Dilioglou, E. F. Vanin and D. A. Vignali (2004). "Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector." Nat Biotechnol 22(5): 589-594.

Tan, P. L., J. E. Wagner, A. D. Auerbach, T. E. Defor, A. Slungaard and M. L. Macmillan (2006). "Successful engraftment without radiation after fludarabine-based regimen in Fanconi anemia patients undergoing genotypically identical donor hematopoietic cell transplantation." Pediatr Blood Cancer 46(5): 630-636.

Timmers, C., T. Taniguchi, J. Hejna, C. Reifsteck, L. Lucas, D. Bruun, M. Thayer, B. Cox, S. Olson, A. D. D'Andrea, R. Moses and M. Grompe (2001). "Positional cloning of a novel Fanconi anemia gene, FANCD2." Mol Cell 7(2): 241-248.

Tolar, J., P. S. Becker, D. W. Clapp, H. Hanenberg, C. D. de Heredia, H. P. Kiem, S. Navarro, P. Qasba, P. Rio, M. Schmidt, J. Sevilla, E. Verhoeyen, A. J. Thrasher and J. Bueren (2012). "Gene therapy for Fanconi anemia: one step closer to the clinic." Hum Gene Ther 23(2): 141-144.

Touzot, F., D. Moshous, R. Creidy, B. Neven, P. Frange, G. Cros, L. Caccavelli, J. Blondeau, A. Magnani, J. M. Luby, B. Ternaux, C. Picard, S. Blanche, A. Fischer, S. Hacein-Bey-Abina and M. Cavazzana (2015). "Faster T-cell development following gene therapy compared with haploidentical HSCT in the treatment of SCID-X1." Blood 125(23): 3563-3569.

Tulpule, A., M. W. Lensch, J. D. Miller, K. Austin, A. D'Andrea, T. M. Schlaeger, A. Shimamura and G. Q. Daley (2010). "Knockdown of Fanconi anemia genes in human embryonic stem cells reveals early developmental defects in the hematopoietic lineage." Blood 115(17): 3453-3462.

Urnov, F. D., J. C. Miller, Y. L. Lee, C. M. Beausejour, J. M. Rock, S. Augustus, A. C. Jamieson, M. H. Porteus, P. D. Gregory and M. C. Holmes (2005). "Highly efficient endogenous human gene correction using designed zinc-finger nucleases." Nature 435(7042): 646-651.

Urnov, F. D., E. J. Rebar, M. C. Holmes, H. S. Zhang and P. D. Gregory (2010). "Genome editing with engineered zinc finger nucleases." Nat Rev Genet 11(9): 636-646.

Valton, J., F. Daboussi, S. Leduc, R. Molina, P. Redondo, R. Macmaster, G. Montoya and P. Duchateau (2012). "5'-Cytosine-phosphoguanine (CpG) methylation impacts the activity of natural and engineered meganucleases." J Biol Chem 287(36): 30139-30150.

Vaz, F., H. Hanenberg, B. Schuster, K. Barker, C. Wiek, V. Erven, K. Neveling, D. Endt, I. Kesterton, F. Autore, F. Fraternali, M. Freund, L. Hartmann, D. Grimwade, R. G. Roberts, H. Schaal, S. Mohammed, N. Rahman, D. Schindler and C. G. Mathew (2010). "Mutation of the RAD51C gene in a Fanconi anemia-like disorder." Nat Genet 42(5): 406-409.

Velazquez, I. and B. P. Alter (2004). "Androgens and liver tumors: Fanconi's anemia and non-Fanconi's conditions." Am J Hematol 77(3): 257-267.

Virts, E. L., A. Jankowska, C. Mackay, M. F. Glaas, C. Wiek, S. L. Kelich, N. Lottmann, F. M. Kennedy, C. Marchal, E. Lehnert, R. E. Scharf, C. Dufour, M. Lanciotti, P. Farruggia, A. Santoro, S. Savasan, K. Scheckenbach, J. Schipper, M. Wagenmann, T. Lewis, M. Leffak, J. L. Farlow, T. M. Foroud, E. Honisch, D. Niederacher, S. C. Chakraborty, G. H. Vance, D. Pruss, K. M. Timms, J. S. Lanchbury, A. F. Alpi and H. Hanenberg (2015). "AluY-mediated germline deletion, duplication and somatic stem cell reversion in UBE2T defines a new subtype of Fanconi anemia." Hum Mol Genet.

Walsh, C. E., K. Fu and M. Brecher (2001). "Retroviral-mediated gene transfer for Fanconi anemia group A patients - a clinical trial." Blood 98.

Wang, A. T., T. Kim, J. E. Wagner, B. A. Conti, F. P. Lach, A. L. Huang, H. Molina, E. M. Sanborn, H. Zierhut, B. K. Cornes, A. Abhyankar, C. Sougnez, S. B. Gabriel, A. D. Auerbach, S. C. Kowalczykowski and A. Smogorzewska (2015). "A Dominant Mutation in Human RAD51 Reveals Its Function in DNA Interstrand Crosslink Repair Independent of Homologous Recombination." Mol Cell 59(3): 478-490.

Wang, H., H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, F. Zhang and R. Jaenisch (2013). "One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering." Cell 153(4): 910-918.

Xia, F., D. G. Taghian, J. S. DeFrank, Z. C. Zeng, H. Willers, G. Iliakis and S. N. Powell (2001). "Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining." Proc Natl Acad Sci U S A 98(15): 8644-8649.

Xie, F., L. Ye, J. C. Chang, A. I. Beyer, J. Wang, M. O. Muench and Y. W. Kan (2014). "Seamless gene correction of beta-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac." Genome Res 24(9): 1526-1533.

Yagasaki, H., D. Adachi, T. Oda, I. Garcia-Higuera, N. Tetteh, A. D. D'Andrea, M. Futaki, S. Asano and T. Yamashita (2001). "A cytoplasmic serine protein kinase binds and may regulate the Fanconi anemia protein FANCA." Blood 98(13): 3650-3657.

Yan, C. and D. D. Boyd (2006). "Histone H3 acetylation and H3 K4 methylation define distinct chromatin regions permissive for transgene expression." Mol Cell Biol 26(17): 6357-6371.

Yang, H., H. Wang, C. S. Shivalila, A. W. Cheng, L. Shi and R. Jaenisch (2013). "One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering." Cell 154(6): 1370-1379.

Yang, X. (2015). "Applications of CRISPR-Cas9 mediated genome engineering." Mil Med Res 2: 11.

Ye, L., J. Wang, A. I. Beyer, F. Teque, T. J. Cradick, Z. Qi, J. C. Chang, G. Bao, M. O. Muench, J. Yu, J. A. Levy and Y. W. Kan (2014). "Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Delta32 mutation confers resistance to HIV infection." Proc Natl Acad Sci U S A 111(26): 9591-9596.

Yung, S. K., K. Tilgner, M. H. Ledran, S. Habibollah, I. Neganova, C. Singhapol, G. Saretzki, M. Stojkovic, L. Armstrong, S. Przyborski and M. Lako (2013). "Brief report: human pluripotent stem cell models of fanconi anemia deficiency reveal an important role for fanconi anemia proteins in cellular reprogramming and survival of hematopoietic progenitors." Stem Cells 31(5): 1022-1029.

Zecca, M., L. Strocchio, D. Pagliara, P. Comoli, A. Bertaina, G. Giorgiani, C. Perotti, F. Corbella, L. Brescia and F. Locatelli (2014). "HLA-haploidentical T cell-depleted allogeneic hematopoietic stem cell transplantation in children with Fanconi anemia." Biol Blood Marrow Transplant 20(4): 571-576.

Zhang, X., D. P. Sejas, Y. Qiu, D. A. Williams and Q. Pang (2007). "Inflammatory ROS promote and cooperate with the Fanconi anemia mutation for hematopoietic senescence." J Cell Sci 120(Pt 9): 1572-1583.

Zhang, X., X. Shang, F. Guo, K. Murphy, M. Kirby, P. Kelly, L. Reeves, F. O. Smith, D. A. Williams, Y. Zheng and Q. Pang (2008). "Defective homing is associated with altered Cdc42 activity in cells from patients with Fanconi anemia group A." Blood 112(5): 1683-1686.

Zou, J., C. L. Sweeney, B. K. Chou, U. Choi, J. Pan, H. Wang, S. N. Dowey, L. Cheng and H. L. Malech (2011). "Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting." Blood 117(21): 5561-5572.

APPENDIX

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

Stem Cell Reports
Article

ISSCR

OPEN ACCESS

Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells

Zita Garate,[1,2,3,12] Oscar Quintana-Bustamante,[1,2,12,*] Ana M. Crane,[3] Emmanuel Olivier,[4] Laurent Poirot,[5] Roman Galetto,[5] Penelope Kosinski,[6] Collin Hill,[6] Charles Kung,[6] Xabi Agirre,[7] Israel Orman,[1,2] Laura Cerrato,[1,2] Omaira Alberquilla,[1,2] Fatima Rodriguez-Fornes,[1,2] Noemi Fusaki,[8] Felix Garcia-Sanchez,[9] Tabita M. Maia,[10] Maria L. Ribeiro,[10] Julian Sevilla,[11] Felipe Prosper,[7] Shengfang Jin,[6] Joanne Mountford,[4] Guillermo Guenechea,[1,2] Agnes Gouble,[5] Juan A. Bueren,[1,2] Brian R. Davis,[3] and Jose C. Segovia[1,2,*]

[1]Hematopoietic Innovative Therapies Division, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Centro de Investigaciones Biomédicas en Red de Enfermedades Raras (CIBERER), Madrid 28040, Spain
[2]Advanced Therapies Mixed Unit, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz (IIS-FJD, UAM), Madrid 28040, Spain
[3]Center for Stem Cell and Regenerative Medicine, Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center, Houston, TX 77030, USA
[4]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow G12 8QQ, UK
[5]Cellectis, Paris 75013, France
[6]Agios Pharmaceuticals, Cambridge, MA 02139-4169, USA
[7]Hematology and Cell Therapy, Clinica Universidad de Navarra and CIMA, Pamplona 31008, Spain
[8]JST PRESTO and Ophthalmology, Keio University, Tokyo 108-8345, Japan
[9]Histocompatibility and Molecular Biology Laboratory, Centro de Transfusion de Madrid, Madrid 28032, Spain
[10]Serviço de Hematologia, Centro Hospitalar e Universitario de Coimbra, Coimbra 3000-075, Portugal
[11]Hospital Universitario Niño Jesús, Madrid 28009, Spain
[12]Co-first author
*Correspondence: oscar.quintana@ciemat.es (O.Q.-B.), jc.segovia@ciemat.es (J.C.S.)
http://dx.doi.org/10.1016/j.stemcr.2015.10.002

SUMMARY

Pyruvate kinase deficiency (PKD) is a rare erythroid metabolic disease caused by mutations in the *PKLR* gene. Erythrocytes from PKD patients show an energetic imbalance causing chronic non-spherocytic hemolytic anemia, as pyruvate kinase defects impair ATP production in erythrocytes. We generated PKD induced pluripotent stem cells (PKDiPSCs) from peripheral blood mononuclear cells (PB-MNCs) of PKD patients by non-integrative Sendai viral vectors. PKDiPSCs were gene edited to integrate a partial codon-optimized R-type pyruvate kinase cDNA in the second intron of the *PKLR* gene by TALEN-mediated homologous recombination (HR). Notably, we found allele specificity of HR led by the presence of a single-nucleotide polymorphism. High numbers of erythroid cells derived from gene-edited PKDiPSCs showed correction of the energetic imbalance, providing an approach to correct metabolic erythroid diseases and demonstrating the practicality of this approach to generate the large cell numbers required for comprehensive biochemical and metabolic erythroid analyses.

INTRODUCTION

Pyruvate kinase deficiency (PKD; OMIM: 266200) is a rare metabolic erythroid disease caused by mutations in the *PKLR* gene, which codes the R-type pyruvate kinase (RPK) in erythrocytes and L-type pyruvate kinase (LPK) in hepatocytes. Pyruvate kinase (PK) catalyzes the last step of glycolysis, the main source of ATP in mature erythrocytes (Zanella et al., 2007). PKD is an autosomal-recessive disease and the most common cause of chronic non-spherocytic hemolytic anemia. The disease becomes clinically relevant when RPK activity decreases below 25% of the normal activity in erythrocytes. PKD treatment is based on supportive measures, such as periodic blood transfusions and splenectomy. The only definitive cure for PKD is allogeneic bone marrow transplantation (Suvatte et al., 1998; Tanphaichitr et al., 2000). However, the low availability of compatible donors and the risks associated with allogeneic bone marrow transplantation limit its clinical application. Transplantation of gene-corrected autologous hematopoietic progenitors might solve these problems. We have developed different gamma-retroviral and lentiviral vectors to correct a mouse PKD model (Meza et al., 2009), and their efficacy is currently being tested in hematopoietic progenitors from PKD patients (M. Garcia-Gomez et al., personal communication). However, the main drawback of current gene therapy approaches based on retro-/lentiviral vectors is the random integration of transgenes, which can promote insertional mutagenesis by disrupting tumor suppressor genes or *cis*-activating proto-oncogenes (Cavazza et al., 2013).

Over the last few years, gene editing by homologous recombination (HR) has been widely used in human cells to avoid undesirable transgene insertion. HR efficacy is very limited in human cells, estimated at one HR event per 10^6 cells; however, the potential application of HR in

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

human cells has been enhanced considerably by sequence-specific DNA nucleases (Carroll, 2011; Porteus and Carroll, 2005). Three different gene-editing strategies can be applied: gene correction, where a mutation is exchanged directly by the wild-type sequence; knockin, where a partial cDNA is inserted in the target locus to express a chimeric mRNA formed by endogenous first exons and partial cDNA under the endogenous promoter control; and safe harbor, in which the transgene is inserted by HR in a safe place in the genome, such as AAVS1 or CCR5 loci (Garate et al., 2013).

Concurrent with the application of gene editing in human cells, the generation of human induced pluripotent stem cells (iPSCs) was described (Takahashi et al., 2007; Yu et al., 2007). iPSCs possess properties of self-renewal and pluripotency that are similar to those of embryonic stem cells (ESCs), but potential alloreactivity and ethical issues associated with human ESCs are avoided. The wide reproducibility of the iPSC technology, independent of cell type and reprogramming methods, has established their great potential for future cell therapies. Additionally, patient- or disease-specific iPSCs are becoming established as in vitro systems to model diseases and to explore new therapeutic approaches. Reprogramming of easily accessible cell sources such as skin fibroblasts (Park et al., 2008), keratinocytes (Aasen et al., 2008), or even peripheral blood mononuclear cells (PB-MNCs) (Loh et al., 2009; Ye et al., 2009) has been described, and many efforts are being made to improve the safety and efficacy of the reprogramming method. Recently, iPSC generation by a Sendai viral vector platform (SeV) (Fusaki et al., 2009; Nishimura et al., 2011), even from blood cells (Nishishita et al., 2011; Seki et al., 2010), has been described as a non-integrative and highly efficient platform.

The correction of patient-specific iPSCs by homologous recombination has been explored in different pathologies (Garate et al., 2013; Karakikes et al., 2015; Rio et al., 2014; Sebastiano et al., 2011; Song et al., 2015), demonstrating its feasibility and setting up gene editing for other stem cells. Herein, we have assessed the combination of cell reprogramming and gene editing for PKD correction as a first example of the possible application of these advanced technologies to metabolic diseases affecting the erythroid lineage. PKD patient-specific iPSCs were efficiently generated from PB-MNCs by an SeV non-integrative system. The *PKLR* gene was edited by PKLR transcription activator-like effector nucleases (TALENs) to introduce a partial codon-optimized cDNA in the second intron by HR. Surprisingly, we found allelic specificity in the HR induced by the presence of a single nucleotide exchange (SNP), demonstrating the potential to select the allele to be corrected. Significantly, a high number of erythroid cells derived from PKDiPSCs was generated and displayed the energetic imbalance characteristic of PKD patients, which was corrected after gene editing.

RESULTS

Generation of Integration-free Specific iPSCs Derived from the Peripheral Blood of PKD Patients

First, to evaluate the potential use of PB-MNCs as a cell source to be reprogrammed to iPSCs by the non-integrative SeV, we analyzed the susceptibility of these cells to SeV. PB-MNCs were expanded in the presence of specific cytokines (stem cell factor [SCF], thrombopoietin [TPO], FLT3L, granulocyte colony-stimulating factor [G-CSF], and IL-3) to promote the maintenance and proliferation of hematopoietic progenitors and myeloid-committed cells for 4 days. Cells were then infected with an SeV encoding for the Azami green fluorescent marker. Five days later, the transduction of hematopoietic progenitor ($CD34^+$), myeloid ($CD14^+$/$CD15^+$), and lymphoid T ($CD3^+$) and B ($CD19^+$) cells was evaluated by flow cytometry. Although the majority of cells in the culture expressed T or B lymphoid markers, a reduced proportion of them (10% of T cells, 3% of B cells) expressed Azami green. In contrast, 54% of the myeloid cells and 76% of the hematopoietic progenitors present in the culture were positive for the fluorescent marker (data not shown), demonstrating that SeV preferentially transduces the less abundant hematopoietic progenitors and myeloid cells under these culture conditions.

This transduction protocol was then used to reprogram PB-MNCs from healthy donors and PKD patients by SeV encoding the four "Yamanaka" reprograming factors (OCT3/4, KLF4, SOX2, and c-MYC; Figure 1A). ESC-like colonies were obtained from one healthy donor (PB2) and from samples from two PKD patients (PKD2 and PKD3) PB-MNCs. Up to 20 ESC-like colonies derived from PB2, 100 from PKD2 and 50 from PKD3 were isolated and expanded (Figure 1B). The complete reprogramming of the different established lines toward embryonic stem (ES)-like cells was evaluated (Figures S1A–S1C). RT-PCR gene expression array verified a similar expression level of the main genes involved in pluripotency and self-renewal in our reprogramed cells and in the reference human ESC line H9 (Figures S1A–S1C). The ES markers *OCT3/4*, *SSEA4*, and *Tra-1-60* were also corroborated by fluorescence-activated cell sorting (FACS) and immunofluorescence (Figures S1A–S1C). Unmethylated status of *NANOG* and *SOX2* promoters was confirmed by pyrosequencing. *NANOG* promoter was strongly demethylated in lines derived from PB2, PKD2, and PKD3. Surprisingly, the *SOX2* promoter was already unmethylated in PB-MNCs (Figure S1D). Furthermore, the pluripotency of these lines derived from PB-MNCs was affirmed by their ability to

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

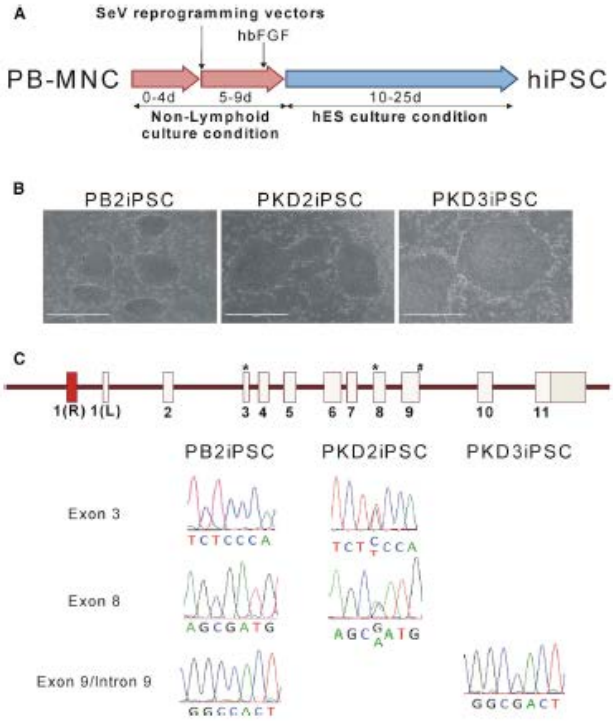

Figure 1. PB-MNC Reprogramming by SeV

PB-MNCs from healthy donors and PKD patients were reprogrammed by SeV expressing OCT4, SOX2, KLF4, and cMYC mRNAs. Several lines from a healthy donor (PB2iPSC), patient PKD2 (PKD2iPSC), and patient PKD3 (PKD3iPSC) were isolated, expanded, and characterized.

(A) Diagram of the reprogramming protocol.

(B) Representative microphotographs of different iPSC lines derived from PB2 MNC, PKD2 MNC, or PKD3 MNC. Scale bars represent 200 μm.

(C) Sanger sequencing of each patient-specific mutation in the *PKLR* gene in PB2iPSC, PKD2iPSC, and PKD3iPSC. *Mutations present in patient PKD2. #Mutation present in patient PKD3.

See also Figures S1 and S2.

generate teratomas into NOD.Cg-PrkdcscidIL2rgtm/Wjl/SzJ (NSG) mice, where all the mice injected developed teratomas showing tissues from the three different embryonic layers (Figures S1A–S1C). These data confirmed the reprogrammed lines as bona fide iPSC lines denoted as PB2iPSC, PKD2iPSC, and PKD3iPSC. Additionally, the presence of the wild-type (WT) sequence or patient specific mutations in the different human iPSC lines generated was confirmed by Sanger sequencing of the corresponding genome loci (Figure 1C). PKD2iPSC showed the two heterozygous mutations in exon 3 (359C > T) and exon 8 (1168G > A), and PKD3iPSC carried the homozygous mutation in the splicing donor sequence of exon 9/intron 9 (IVS9(+1)G > C) characterized in the patients. These mutations could not be detected in ~~peripheral-blood-derived induced pluripotent stem cells (PBiPSCs),~~ which showed the expected WT sequences (Figures 1C).

To confirm the absence of ectopic reprogramming gene expression, we analyzed the disappearance of SeV vectors in the generated iPSCs. The presence of the ectopic proteins could be tracked by the persistence of the fluorescent marker, as the SeV expressing Azami green was co-transduced together with the reprogramming vectors. Azami green expression was only detected in non-reprogramed, fibroblast-like cells in early passages. Green fluorescence disappeared in all the iPSC colonies (Figure S1E). Importantly, SeV mRNA was not detected in iPSCs derived from PB-MNCs in late passages (Figure S1E).

In addition, to check whether the established protocol did allow preferential reprogramming in myeloid and/or progenitor cells, T cell receptor (TCR) and immunoglobulin heavy-chain genome rearrangements were studied on the iPSC generated (Figure S2). None of the analyzed iPSC clones (PB2iPSC c33, PKD2iPSC c78, PKD3iPSC c14, PKD3iPSC c10, and PKD3iPSC c35) had any T or B rearrangements, meaning that iPSC clones were generated from neither T nor B lymphocytes. These results guarantee the SeV-based reprograming system as the best option in

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

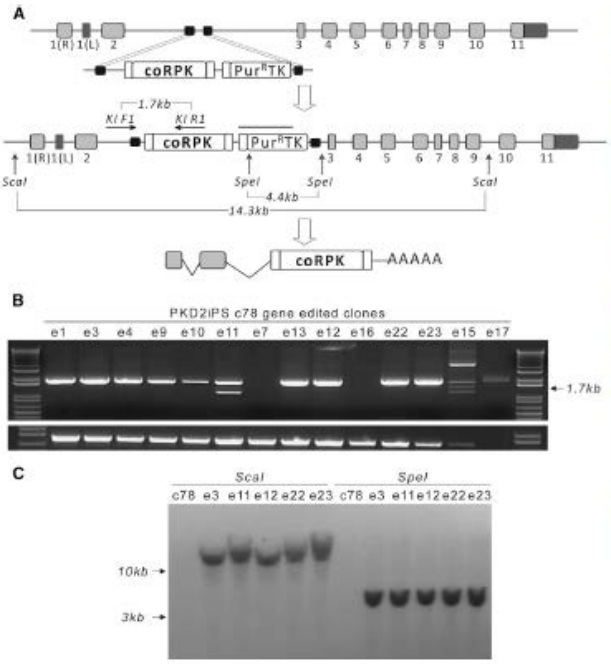

Figure 2. Gene Editing in the PKLR Locus

(A) Diagram showing where therapeutic matrix is introduced by HR in the *PKLR* locus. The strategy to identify the integrated matrix by PCR (horizontal arrows) and Southern blot (vertical arrows) indicating the expected DNA fragment sizes is shown, and the line over the Puro^R/thymidine kinase fusion cassette indicates probe location. Small squares at the beginning and end of the partial codon-optimized (cDNA) *RPK* indicate splicing acceptor and FLAG tag sequences present in the cassette, respectively; light gray squares represent endogenous (mRNA)*RPK* exons; dark gray squares represent the first LPK exon and 3′ UTRs at the beginning and at the end of the PKLR gene, respectively; and black squares represent homology arms.

(B) DNA electrophoresis of gDNA from Puro^R-PKD2iPSC clones, amplified by PCR to identify specific matrix integration.

(C) Southern blot of gDNA from edited PKD2iPSC clones, digested by ScaI or SpeI to confirm the precise integration of the matrix in the *PKLR* locus.

See also Figures S3 and S4 and Table S2.

reprogramming peripheral blood, as the reprograming vectors are cleared after iPSC generation, and the iPSC are generated from non-lymphoid cells. To continue with the following gene-editing steps clones from PB2, PKD2, and PKD3, we randomly selected ~~PB-MNCs.~~

TALEN-Based Gene Editing in the PKLR Locus of PKDiPSCs

To achieve correction of PKDiPSCs, we used a knockin gene-editing strategy based on inserting a therapeutic matrix containing a partial codon-optimized (cDNA)*RPK* gene covering exons 3 to 11, fused to a FLAG tag and preceded by a splice acceptor signal. Additionally, a positive-negative selection cassette containing a puromycin (Puro) resistance/thymidine kinase (TK) fusion gene driven by mouse phosphoglycerate kinase (mPGK) promoter was included downstream of the partial (cDNA)*RPK*. These elements were flanked by two homology arms matching sequences in the second intron of the *PKLR* gene (Figure 2A). In order to increase the efficiency of gene editing, we developed a *PKLR*-specific TALEN targeting a specific genomic sequence in the second intron flanked by the homology arms. Nuclease activity of the PKLR TALEN in the target sequence was verified by surveyor assay after nucleofecting both subunits of the nuclease in PKD2iPSC and PKD3iPSC (data not shown).

In two independent experiments, two iPSC lines from two different PKD patients, PKD2iPSC c78 and PKD3iPSC c54, were nucleofected with a control plasmid or with the developed matrix (from now on called therapeutic matrix) alone or together with two different doses of PKLR TALEN (1.5 or 5 μg of each PKLR TALEN subunit). Two days later, Puro was added to the media for 1 week. Puro-resistant (Puro^R) colonies, with a satisfactory morphology appeared and were individually picked and subcloned. Most of the Puro^R colonies were identified from cells nucleofected with both the matrix and the PKLR TALEN subunits, although some colonies grew out after receiving only the therapeutic matrix. There was no difference in the number of Puro^R colonies between PKDiPSC lines from the different patients. To confirm target insertion of the therapeutic matrix in the second intron of the *PKLR* gene, we performed specific PCR analyses (Figures 2A and S3). The expected PCR product was detected in 10 out of 14 Puro^R clones

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

Table 1. Efficacy of Homologous Recombination in PKD2iPSCs and PKD3iPSCs and Indels Analysis in the Untargeted Allele

	Puro^R Clones	Percentage of Gene-Edited Clones	Percentage of Gene-Edited Clones Targeted Biallelically	Percentage of Gene-Edited Clones with Indels in the Untargeted Allele
PKD2iPSCs	13	77%	0%	40%
PKD3iPSCs	40	76%	11%	31%

from PKD2iPSC c78 and 31 out of 40 Puro^R clones from PKD3iPSC c54 (Figures 2B and S3). Taken together, we estimated an HR frequency among the Puro^R clones of above 75% for the two reprogramed patients (Table 1). In addition, two Puro^R clones from PKD3iPSC c54 clone nucleofected with the therapeutic matrix alone were positive for knockin, estimating an efficiency of 0.6 edited per 1×10^5 nucleofected cells. Despite detecting HR without nucleases, the HR frequency was boosted almost five times (2.85 edited PKD3iPSC per 1×10^5 nucleofected cells) when the PKLR TALEN was added. Additionally, knockin insertion of the therapeutic matrix was verified by Southern blot (Figure 2C), confirming a single insertion in the desired genomic locus.

Next, we tested whether the PKLR TALEN was also cutting the untargeted allele. Up to 40% of PKD2 and 31% of PKD3 edited clones carried insertions-deletions (indels) in the untargeted allele of the PKLR TALEN target site (Table 1; Figure S3C), demonstrating the high efficacy of this PKLR TALEN. Moreover, 3 out of 40 edited clones from PKD3iPSC were targeted biallelically as determined when both the targeted allele and the untargeted were analyzed in a single PCR (Figure S3D). In contrast, no edited PKD2iPSC clones showed biallelic targeting.

In order to check the specificity of the PKLR TALEN, we looked for potential off-target cutting sites in the different edited PKDiPSC clones. By in silico studies, we found five hypothetical off-target sites for this TALEN (Table S2). These five off-targets can be recognized by the two subunits matched as homodimers or heterodimer, where the left subunit can join the right subunit or each subunit can join a different spacer sequence and length (Table S2). All the potential off-targets had at least five mismatched bases, which makes the recognition by the TALEN unlikely. To confirm the specificity of the TALEN, we amplified genomic DNA from several edited PKD2iPSC and PKD3iPSC clones and Sanger sequenced around four off-targets (off-targets 1, 2, 4, and 5; data not shown). None of the analyzed clones showed any indels in any of the off-targets analyzed. Off-target 3 could not be amplified by PCR. Nevertheless, as the first base in the 5′ recognition site of the off-target 3 was an A, the recognition of this off-target by the PKLR TALEN is strongly reduced (Boch et al., 2009). This high specificity together with the high efficacy of PKLR TALEN confirms the feasibility of the developed TALEN and therapeutic matrix to promote HR in the *PKLR* locus.

Finally, we verified the pluripotency of the edited iPSCs after gene editing by in vivo teratoma formation into NSG mice (Figure S4). Edited clones were able to generate teratomas with tissues from the three embryonic layers. More importantly, human hematopoiesis, demonstrated by the presence of cells expressing the human CD45 panleukocytary marker (4.54% of the total teratoma forming cells) and human progenitors (CD45$^+$CD34$^+$; 2.74% of the total hCD45$^+$ cells) derived from edited PKD3iPSC e31 teratomas could also be detected in vivo (Figure S4B). Altogether, the data confirm the use of PKLR TALEN to edit the *PKLR* gene in PKDiPSCs without affecting their pluripotent properties.

A Single-Nucleotide Polymorphism Leads to Allele-Specific Targeting

While evaluating the presence of indels in the untargeted allele by Sanger sequencing, we identified the existence of a g.[2268A > G] SNP 43 bases apart from the PKLR TALEN cutting site in PKD2iPSC (Figure 3A). Interestingly, the untargeted allele from all the edited PKD2iPSC clones (ten out of ten) carried the previously mentioned SNP, suggesting an impediment of the allele carrying the SNP variant to carry out HR. Moreover, no biallelic targeting was detected in any PKD2iPSC edited clone. On the contrary, 3 out of 31 edited PKD3iPSC clones without any SNP in the homology genomic area were targeted in both alleles.

Genetic Stability of PKDiPSCs and Gene-Edited PKDiPSCs

We wanted to study whether the whole process of reprogramming plus gene editing was inducing genetic instability in the resulting cells. As a first approach, we performed karyotyping of the different iPSC lines and confirmed normal karyotype in all cases (data not shown). However, to have a clearer assessment, we monitored the genetic stability throughout all the process, including iPSC generation and gene-editing correction, by comparative genomic hybridization (CGH) and exome sequencing. PB-MNCs from a PKD2 patient, reprogrammed PKD2iPSC c58, and edited PKD2iPSC e11 were selected as representatives of each step. Copy-number variations (CNVs) were defined in these samples after comparing with a reference genomic

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

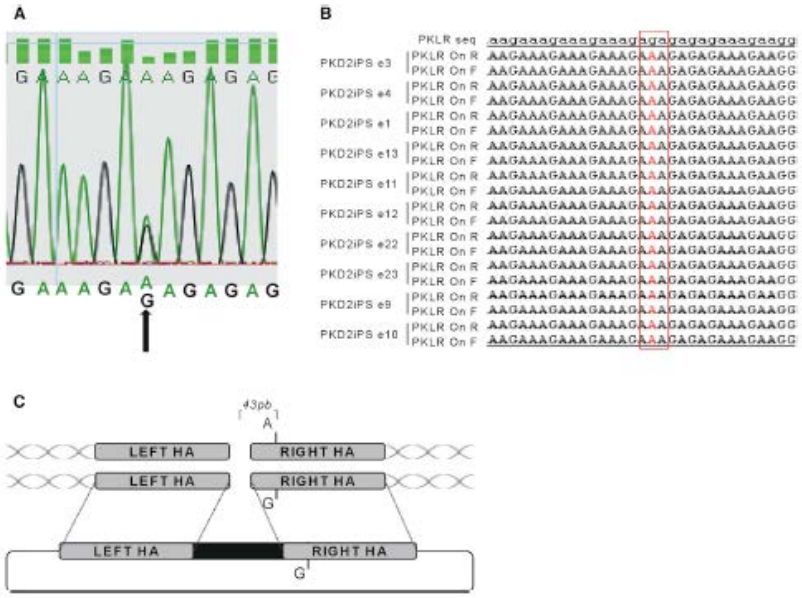

Figure 3. Allele-Specific Targeting on the PKLR Locus

(A) A single-nucleotide polymorphism (SNP) detected in the second intron of the PKLR gene in PKD2 patient cells, identified by Sanger sequencing. Black arrow points to the polymorphism.

(B) Sequence of PKD2 SNP in the untargeted allele in all the edited PKD2iPSC clones. Letter in red indicates the SNP.

(C) Diagram indicating the position of the SNP with respect to the theoretical cutting site of the PKLR TALEN and the matrix integration in the targeted allele.

DNA. Among the total CNVs identified, 31 were present in the original PB-MNC from PKD2, 34 CNVs were detected in PKD2iPSC c78, and 32 in PKD2iPSC e11 (Table 2; Table S3). Twenty-three CNVs detected in PKD2iPSC c78 were already present in PKD2 PB-MNCs, indicating the mosaicism of the original patient sample. On the other hand, only four CNVs present in PKD2iPSC c78 and PKD2iPSC e11 were not detected in the primary sample. Of note, these four CNV were at chromosomes 1q44, 2p21, 3p12.3-p12.1, and Xp11.22, involving genes such as *ROBO1*, *GBE1*, *TCEA1*, *LYPLA1*, *DLG2*, *PLEKHA5*, and *AEBP2* (Table 2). More importantly, only two CNVs appeared after gene editing that were not present in the original iPSC clone. The first one was a deletion of 6.6 kb that include several olfactory receptor genes (such as *OR2T11*, *OR2T35*, or *OR2T27*), and the second CNV was an amplification of 0.6 kb that includes the *FGD1* gene. Additionally, sequences surrounding these two CNVs in PKD2iPSC e11 have more than eight mismatches with the PKLR TALEN recognition site, suggesting that these genomic alterations were not produced by gene editing. Moreover, we analyzed the presence of CNVs in PKD3iPSC before and after gene editing to confirm the potential harmless effect in the genomic stability of PKLR TALEN activity (Table S4). Edited clone PKD3iPSC e31 (biallelically targeted) showed 10 out 11 CNVs of the parental PKD3iPSC c54, and PKD3iPSC e88 (monoallelically targeted) showed two new CNVs. Furthermore, none of the CNVs present in the edited PKD2iPSC e11 were present in any of these two PKD3iPSC edited clones, which suggests that PKLR TALEN does not induce any specific CNVs in PKDiPSC clones.

Simultaneously, the three PKD2 samples were assayed using the Illumina HiSeq 2000 system for exome sequencing. After bioinformatics analysis by comparing the sequencing data with a human genome reference, PKD2 PB-MNCs showed 68,260 changes in their sequences, PKD2iPSC c78 68,542, and PKD2iPSC e11 67,728 (Table S5). Only ten of all variants detected in PKD2iPSC e11 were in exonic regions, included in the SNP database, and not identified in

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

Table 2. Copy-Number Variations and Exome Variants Detected by CGH and Exome Sequencing in Edited PKD2iPSCs

CGH Analysis

Number	Chromosome	Cytoband	Size (bp)	Type	Present in PKD2iPSC c78	Present in PKD2 PB-MNCs
1	1	q44	60,641	DEL	no	no
2	3	p12.2-p12.1	3,931,633	LOH	yes	no
3	8	q11.23	169,460	AMP	yes	no
4	11	q14.1	113,264	DEL	yes	no
5	12	p12.3	1,182,747	AMP	yes	no
6	17	q21.31	199,747	AMP	yes	no
7	X	p11.22	6,030	AMP	no	no

Exome Sequencing

Number	Chromosome	Reference Base	Altered Base	Gene	Type	Present in PKD2iPSC c78
1	9	–	TGCCTCCACCACACC	PHF2	nonframeshift insertion	no
2	16	G	T	ZNF747	nonsynonymous SNV	no
3	6	G	C	SNX3	nonsynonymous SNV	no
4	22	A	T	TUBGCP6	nonsynonymous SNV	no
5	10	A	G	TARC2	nonsynonymous SNV	no
6	7	C	A	TNRC18	stop-gain SNV	no
7	18	C	A	MBD2	nonsynonymous SNV	yes
8	18	C	A	MBD2	nonsynonymous SNV	yes
9	9	G	T	RUSC2	nonsynonymous SNV	yes
10	11	G	A	APOA5	nonsynonymous SNV	yes

SNV, single-nucleotide variation.
See also Tables S4 and S5.

PKD2 PB-MNCs (Table 2). Additionally, four of them were also detected in PKD2iPSC c78. In order to verify the presence of these mutations by Sanger sequencing, we PCR amplified and sequenced these regions. Only the mutations in the *RUSC2*, *TACR2*, and in *APOA5* genes could be confirmed by sequencing (data not shown). None of the ten variants were included in the COSMIC database (Wellcome Trust Sanger Institute, 2014), which includes all the known somatic mutations involved in cancer.

Overall, genetic stability analysis confirmed the safety of our gene editing approach. All the genetic alterations identified were present in the PB-MNCs or generated during their reprogramming or iPSC expansion. Moreover, none of the confirmed alterations could be associated with potentially dangerous mutations.

Gene-Edited PKDiPSCs Recover RPK Functionality

Once the knockin integration was confirmed, we assessed the PK phenotypic correction of the gene-edited iPSCs. We induced the erythroid differentiation of different iPSC lines from a healthy donor iPSC line (PB2iPSC c33), PK-deficient iPSC lines derived from both patients (PKD2iPSC c78 and PKD3iPSC c54), and the corresponding edited clones (monoallelically edited PKD2iPSC e11 and PKD3iPSC e88 and a biallelically targeted PKD3iPSC e31). Characteristic hematopoietic progenitor markers, such as CD43, CD34, and CD45, started to appear over time (data not shown) and were expressed in a similar proportion of cells. Erythroid cells were clearly observed in the cultures (Figure S5A), and the specific erythroid combination of CD71 and CD235a antigens was expressed on the majority of cells after 21 days of differentiation (Figures 4A and S5B). Moreover, cells derived from all iPSC lines analyzed at day 31 of differentiation, showed a similar globin pattern, in which α- and γ-globins were predominant with a small amount of β-globin, and residual embryonic ε- and ζ-globins detected, confirming the erythroid differentiation of these pluripotent lines (Figure S6A). More importantly,

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

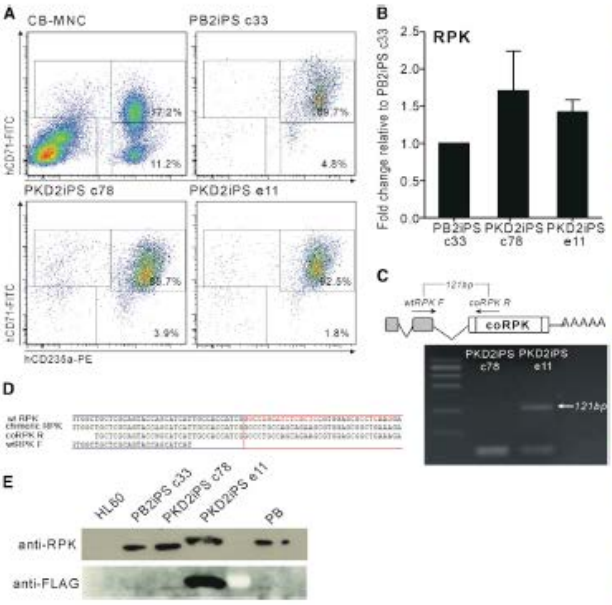

Figure 4. Erythroid Differentiation of PKD2iPSCs

PB2iPSCs, PKD2iPSCs, and edited PKD2iPSCs were differentiated to erythroid cells under specific conditions and analyzed after 31 days in in vitro proliferation and differentiation conditions.

(A) Erythroid differentiation was confirmed by flow cytometry analysis. Cord blood MNCs, PB2iPSC clone c33, PKD2iPC clone c78, and edited PKD2iPSC clone e11 representative analyses are shown.

(B) RPK expression in erythroid cells derived from the different iPSCs was evaluated by qRT-PCR (n = 6).

(C) Specific RT-PCR to amplify the chimeric (mRNA)*RPK* in edited PKD2iPSC. The primers amplified the region around the link between endogenous (mRNA)*RPK* and the introduced codon-optimized (cDNA)*RPK* sequence. Arrow indicates the expected band and the corresponding size only preset in the RNA from edited cells (PKD2iPSC e11).

(D) The sequence of the chimeric transcript was aligned with the theoretical expected sequence after the correct splicing between the endogenous exon 2 (blue square) and the exogenous exon 3 (red square).

(E) The presences of RPK protein in erythroid cells derived from PB2iPSCs, PKD2iPSCs, and edited PKD2iPSCs assessed by western blot (upper line); mobility change in PKD2iPSC e11 is due to the FLAG tag added to the chimeric protein. Expression of chimeric protein was detected by anti-FLAG antibody only in erythroid cells derived from edited PKD2iPSCs (bottom line).

See also Figures S5 and S6.

the erythroid cells derived from the three iPSC lines were able to express RPK (Figures 4B, 4E, S5C, and S5F). It is noteworthy that no alteration in the expression of proximal genes in the edited erythroid cells was confirmed by qRT-PCR (Figure S6B).

The presence of chimeric transcripts in all of the edited PKDiPSC lines was confirmed by RT-PCR. Primers recognizing a sequence in the second endogenous exon of the *PKLR* gene and in the partial codon-optimized (cDNA)*RPK* were able to produce an amplicon with the correct size, specifically in erythroid cells derived from gene-edited PKDiPSCs (Figures 4C and S5E). This amplicon was sequenced and the joint between both parts of the mRNA, coming from the transcription of the endogenous and the exogenous sequences, was detected (Figure 4D). Additionally, the presence of RPK was demonstrated by western blot in the erythroid cells derived from all of the edited iPSC lines derived from PKD2iPSC c78 (PKD2iPSC e11; Figure 4E) and from PKD3iPSC c54 (PKD3iPSC e88 and PKD3iPSC e31; Figure S5E). Interestingly, although (mRNA)*RPK* could be detected in erythroid cells derived from all the iPSC lines derived from PKD3 (Figure S5C), RPK protein was not detected in PKD3iPSC c54 (Figure S6F), probably due to the severity of the mutation in terms of RNA translation. However, the gene edition of PKD3iPSC restored RPK protein expression either in the bialellic (PKD3iPSC e31) and monoallelic (PKD3iPSC e88) edited lines (Figure S5F). Moreover, both the level of the chimeric transcript and the RPK protein were higher in the biallelically targeted clone PKD3iPSC e31 than in the monoallelic PKD3iPSC e88 (Figures S5D and S5F). It is worth it mentioning that flagged RPK was detected in erythroid cells generated after gene editing of PKDiPSCs (Figure 4E), confirming the origin of the RPK protein from the edited genome.

Finally, the recovery in metabolic function of the corrected cells was assessed in the differentiated cells by conventional biochemical analysis as well as by liquid chromatography mass spectrometry (LC-MS) (Figures 5 and S7). The ATP level in erythroid cells derived from the monoallelically edited PKDiPSCs (PKD2iPSC e11 and PKD3iPSC

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

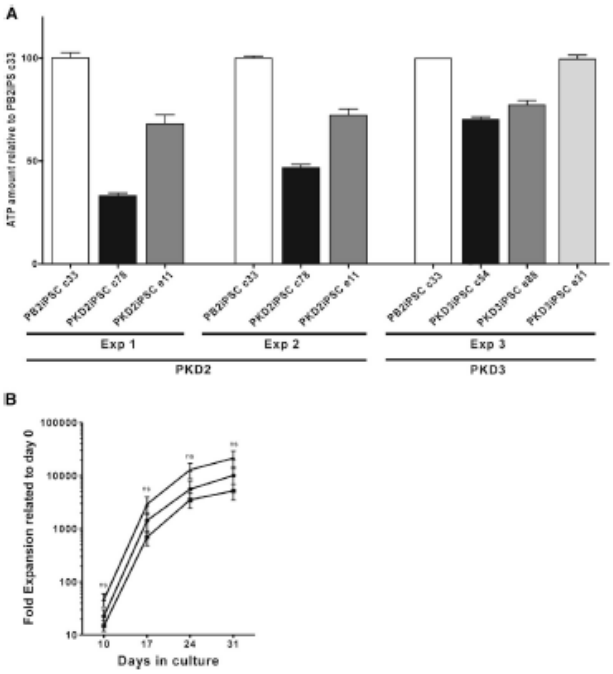

Figure 5. Phenotypic Correction in Edited PKD2iPSCs

(A) ATP levels in erythroid cells derived from healthy iPSCs (PB2iPSCs), PKDiPSCs (patients PKD2 and PKD3), and edited PKDiPSCs (PKD2iPSC e11, PKD3iPSC e88, and PKD3iPSC e31 clones). Data were obtained from three independent experiments from six different iPSC lines derived from two different patients.

(B) In vitro proliferation and differentiation of PB2iPSC clone c33 (■), PKD2iPC clone c78 (▲), and edited PKD2iPSC clone e11 (●). ns, statistically not significant.

See also Figure S7.

e88) was augmented after gene editing (Figure 5A), reaching an intermediate level between that observed in erythroid cells from WT iPSCs and their respective patient-specific iPSC lines. Additionally, erythroid cells derived from the biallelically targeted PKD3iPSC e31 restored the ATP level completely up to healthy values (Figure 5A). In edited erythroid cells, other glycolytic metabolites, such as 2,3-diphosphoglyceric acid, 2-phosphoglyceric acid, pyruvic acid, and L-lactic acid, reached levels between those of control and deficient erythroid cells derived from PB2iPSCs and PKDiPSCs (Figure S7). In addition, we obtained up to 2×10^4-fold expansion of cells in 1 month, meaning that up to 20,000 erythroid cells could be generated from a single iPSC (Figure 5B). As expected, no statistical differences were observed between the different iPSCs, indicating that RPK deficiency only affects the last steps of the erythroid differentiation, where no proliferation is taking place. Altogether, our data validate the effectiveness of this knockin approach to express a corrected RPK protein and demonstrate its potential to therapeutically correct the PKD phenotype and generate large numbers (10^9–10^{10}) of differentiating cells required for comprehensive biochemical and metabolic analyses during their maturation, or even for a potential therapeutic use.

DISCUSSION

In this work, we have shown the potential to combine cell reprograming and gene editing as a therapeutic approach for PKD patients. We generated iPSCs from PB-MNCs taken from PKD patients using a non-integrating viral system. These PKDiPSC lines were effectively gene edited via a knockin strategy at the *PKLR* locus, facilitated by specific PKLR TALENs. More importantly, we have demonstrated the rescue of the disease phenotype in erythroid cells derived from edited PKDiPSCs by the partial restoration of the step of the glycolysis affected in PKD and the improvement of the total ATP level in the erythroid cells derived from PKDiPSCs. The restoration of the energetic balance in erythroid cells derived from PKD patients opens

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

up the possibility of using gene editing to treat PKD patients.

To reprogram patient cells, we adopted the most feasible and safest protocol using a patient cell source that is easy to obtain, PB-MNCs, and an integration-free reprogramming strategy based on SeV vectors. PB-MNCs were chosen, as blood collection is common in patient follow-up and is minimally invasive. Additionally, it is possible to recover enough PB-MNCs from a routine blood collection to perform several reprogramming experiments. Finally, previous works showed that PB-MNCs could be reprogrammed, although at a very low efficiency (Staerk et al., 2010). On the other hand, the SeV reprogramming platform has been described as a very effective, non-integrative system for iPSC reprogramming with a wide tropism for the target cells (Ban et al., 2011; Fusaki et al., 2009). Reprogrammed SeVs are cleared after cell reprogramming due to the difference of replication between newly generated iPSCs and viral mRNA (Ban et al., 2011; Fusaki et al., 2009). However, reprogrammed T or B cells might be favored when whole PB-MNCs are chosen, as these are the most abundant nucleated cell type in these samples. Reprogramming T or B cells has the risk of generating iPSCs with either TCR or immunoglobulin rearrangements, decreasing the immunological repertoire of the hematopoietic cells derived from these rearranged iPSCs. In order to avoid this possibility, we have biased the protocol against reprogramming of either T or B lymphocytes by culturing PB-MNCs with essential cytokines to favor the maintenance and proliferation of hematopoietic progenitors and myeloid cells, as previously shown for retroviral reprogramming vectors (Staerk et al., 2010). This approach was supported here by the demonstration that SeV vectors preferentially transduced hematopoietic progenitors and myeloid cells under these specific conditions and consequently none of the iPSC lines analyzed had immunoglobulin or TCR rearrangements. We further demonstrated that the generation of iPSCs from PB-MNCs using SeV is feasible and simple and generates integration-free iPSC lines with all the characteristic features of true iPSCs that could be further used for research or clinical purposes.

The next goal for gene therapy is the directed insertion of the therapeutic sequences in the cell genome (Garate et al., 2013; Genovese et al., 2014; Karakikes et al., 2015; Song et al., 2015). A number of different gene-editing strategies have been described, including gene modification of the specific mutation, integration of the therapeutic sequences in a safe harbor site, or knockin into the same gene locus. We directed a knockin strategy to insert the partial cDNA of a codon-optimized version of *RPK* in the second intron of the PKLR gene. If used clinically, this strategy would allow the treatment of up to 95% of the patients, those with mutations from the third exon to the end of the (cDNA)*RPK* (Beutler and Gelbart, 2000; Fermo et al., 2005; Zanella et al., 2005). Additionally, this approach retained the endogenous regulation of RPK after gene editing, a necessary factor as RPK is tightly regulated throughout the erythroid differentiation. This fine control would be lost if a safe-harbor strategy was chosen.

The PKLR TALEN generated was very specific and very efficient. We did not find any mutation in any of the theoretical off-target sites defined by the off-site search algorithm and analyzed by PCR and gene sequenced. Moreover, we determined that 2.85 out to 100,000 electroporated PKDiPSCs, without considering the toxicity associated to nucleofection, were gene edited when the PKLR TALEN was used, reaching values similar to those previously published by others (Porteus and Carroll, 2005). Interestingly, 40% of the edited PKDiPSC clones presented indels in the untargeted allele or were biallelically targeted, which indicated that the developed TALEN are very efficient, cutting on the on-target sequence with a high frequency.

Surprisingly, we found that the presence of a single SNP 43 bp away from the PKLR TALEN cutting site was an impediment to HR. The presence of an SNP, which disrupts the complete matching between the genome sequence and homology arm, has already been reported to reduce the frequency of HR (Deyle et al., 2014). Taking into account that the TALEN cut has occurred, as we can detect indels in the non-targeted allele, the absence of matrix insertion seems to be directly related to problems related with the perfect annealing of the matrix with the genome sequences. We have to point out that this SNP is located in a very repetitive region, which might form a structural configuration that increases the HR specificity between this region and its homology arm, as has already been mentioned (Renkawitz et al., 2014). Thus, the genome context where the HR has to take place plays an important role and can facilitate or impair HR. In any case, these data demonstrate the important need for gene-editing strategies to generate the homology arms of an HR matrix from the individual DNA that will be edited. This would restrict HR matrices to patients with similar SNPs in the genomic region to be edited. Therefore, any gene-editing therapy using a knockin or safe-harbor strategy should first screen each patient for the presence of an SNP in the homology arms selected. On the other hand, the presence of a specific SNP could also help to perform allele-specific gene targeting in the cases where the presence of a dominant allele is pathogenic as, for example, in α-thalassemia (De Gobbi et al., 2006).

The gene-editing strategy utilized here to correct PKD was safe, since neither the introduction of genomic alterations nor alteration of the expression of neighboring genes by the insertion and expression of the exogenous sequences occurred. This demonstrates the safety of this knockin gene-editing strategy without *cis* activation of

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

any gene, in comparison to previous results where the selection cassette deregulated nearby genes (Zou et al., 2011). Furthermore, we did not observe any off-target effects induced by PKLR TALEN gene editing.

We found several genomic alterations by CGH and exome sequencing analysis. However, the majority of them were already present in PKD PB-MNCs before their reprogramming, especially in the case of the biallelic targeted PKD3iPSC c31 (Table S4), where all of the CNVs were already present in PKD3iPSC c54, confirming previous data associating these DNA variations in iPSC clones with a cellular mosaicism in the original samples (Abyzov et al., 2012). However, there were some mutations present in the iPSC that we were unable to detect in the original sample, which might be due to technical limitations or to the inherent genetic instability associated with the reprogramming process and iPSC culture (Gore et al., 2011; Hussein et al., 2011). Supporting this last possibility, we found CNVs present in PKD2iPSC c78 and not in PKD2iPSC e11 (Table 2; Table S3). Because PKD2iPSC c78 was maintained in vitro for several more passages, after HR and before CGH analysis, some new changes could have occurred that were not present in the gene-edited-derived clones. Although one CNV involved the *TCEA1* gene, indirectly involved in salivary adenoma as a translocation partner of *PLAG1* (Asp et al., 2006), none of these genomic alterations identified were implicated in hematopoietic malignancies, cell proliferation, or apoptosis regulation, suggesting their neutrality in the PKD therapy by gene editing.

Constitutive expression of Puro/TK from the ubiquitously active mPGK promoter might hinder therapeutic applications of this approach. Indeed, these highly immunogenic prokaryotic/viral proteins can be presented on the cell surface of the gene-corrected cells by the major histocompatibility complex class I molecules, thus stimulating an immune response against the cells once transplanted into the patients. Here, although the Puro/TK cassette has been maintained in the edited PKDiPSC lines, the cassette is inserted between two *loxP* sites, which would allow us to excise it before their clinical application. Moreover, for the potential clinical use of our approach, other selection systems could be used, such as a truncated version of the nerve growth factor receptor combined with enrichment by magnetic sorting, or the use of an inducible or an embryonic-specific promoter instead of the PGK constitutive promoter to limit the Puro/TK expression.

Finally, we have clearly demonstrated the effectiveness of editing the *PKLR* gene in PKDiPSCs to recover the energetic balance in erythroid cells derived from edited PKDiPSCs. ATP and other metabolites involved in glycolysis were restored by expressing a chimeric RPK in a physiological manner. As expected erythroid cells derived from monoallelic corrected PKDiPSCs produce partial restoration of ATP levels, and erythroid cells derived from biallelic corrected PKD3iPSC e31 fully recovered ATP level (Figure 5A). Additionally, we could not observe any difference in the erythroid populations obtained in vitro from uncorrected and corrected PKDiPSCs, probably due to the lack of terminal differentiation/enucleation of the protocol used to generate mature enucleated erythrocytes. Furthermore, we were able to generate 20,000 erythroid cells per starting iPSC, providing abundant material for our assays and offering the potential to undertake more comprehensive analyses, including metabolic and biochemical profiling, to further elucidate the effects of PKD on erythroid cells, or even for therapeutic usage.

Many groups are working to generate long-term reconstituting HSCs from iPSCs, and a major development was reported by Amabile et al. (2013), who showed that in vivo differentiation of human iPSCs in NSG mice reveals their intrinsic potential to fully reconstitute the hematopoietic system. We confirmed the in vivo hematopoietic potential of gene-edited PKDiPSCs (even hematopoietic progenitors could be detected), but we failed to generate in vivo engraftable hematopoietic progenitors (data no shown), possibly because of the low efficacy of our in vivo hematopoietic differentiation approach, which we are working to improve.

In summary, we combined gene editing and patient-specific iPSCs to correct PKD. Our gene-editing strategy was based on inserting a partial codon-optimized (cDNA)*RPK* in the *PKLR* locus mediated by PKLR TALEN without altering the cellular genome or neighbor gene expression. Additionally, we found a highly homologous sequence specificity, since a single SNP could avoid HR. The resultant edited PKDiPSC lines could be differentiated to large number of erythroid cells, where the energetic defect of PKD erythrocytes was effectively corrected. This validates the use of iPSCs for disease modeling and demonstrates the potential future use of gene editing to correct PKD and also other metabolic red blood cell diseases in which a continuous source of fully functional erythrocytes is required.

EXPERIMENTAL PROCEDURES

Peripheral Blood Samples and Reprogramming

Peripheral blood from PKD patients and healthy donors was collected in routine blood sampling from Hospital Clínico Infantil Universitario Niño Jesús (Madrid, Spain), Centro Hospitalario de Coimbra (Coimbra, Portugal), and the Medical Care Service of CIEMAT (Madrid, Spain). All samples were collected under written consent and institutional review board agreement. PB-MNCs were isolated by density gradient using Ficoll-Paque (GE Healthcare). PB-MNCs were pre-stimulated for 4 days in StemSpan (STEMCELL Technologies) plus 100 ng/ml human stem cell factor (SCF), 100 ng/ml hFLT3L, 20 ng/ml hTPO, 10 ng/ml G-CSF, and 2 ng/ml human IL-3 (Peprotech) (Figure 1A). Cells were then

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

transduced with a mix of SeV, kindly provided by DNAvec (Japan), expressing OCT3/4, KLF4, SOX2, c-MYC, and Azami Green, each at a MOI of 3. Transduced cells were maintained for four more days in the same culture medium and then supplemented with 10 ng/ml basic fibroblast growth factor (FGF). Five days after transduction, cells were collected and seeded on irradiated human foreskin fibroblast (HFF-1)-coated (ATCC) culture plates with human ES media (knockout DMEM, 20% knockout serum replacement, 1 mM L-glutamine, and 1% nonessential amino acids [all from Life Technologies]), 0.1 mM β-mercaptoethanol (Sigma-Aldrich), and 10 ng/ml basic human FGF (Peprotech). Human ES media was changed every other day. When human ES-like colonies appeared, they were selected under the stereoscope (Olympus) and a clonal culture from each colony was established.

Gene Editing in iPSCs

iPSCs were treated with Rock inhibitor Y-27632 (Sigma) before a single-cell suspension of iPSCs was generated by StemPro Accutase (Life Technologies) treatment and then nucleofected with 1.5 μg or 5 μg of each PKLR TALEN subunit with or without 4 μg HR matrix by Amaxa Nucleofector (Lonza) using the A23 program. After nucleofection, cells were seeded into a feeder of irradiated Puro^R mouse embryonic fibroblasts in the presence of Y-27632, and 48 hr after transfection, puromycin (0.5 μg/ml) was added to human ES media. Newly formed Puro^R-PKDiPSC colonies were picked individually during a puromycin selection period of 6–10 days. Puro^R-PKDiPSC colonies were expanded and analyzed by PCR and Southern blot to detect HR (Figures 2B, 2C, S3B, and S3D).

Erythroid Differentiation

Erythroid differentiation from iPSC lines was performed using a patented method (WO/2014/013255). In brief, we used a multistep, feeder-free protocol developed by E.O. (unpublished data). Before differentiation, normal, diseased, and corrected iPSCs were maintained in StemPro medium (Life Technologies) with the addition of 20 ng/ml basic FGF on a matrix of recombinant vitronectin fragments (Life Technologies) using manual passage. For initiation of differentiation, embryoid bodies (EBs) were formed in Stemline II medium (Sigma Aldrich) with BMP4, vascular endothelial growth factor (VEGF), Wnt3a, and activin A. In a second step, hematopoietic differentiation was induced by adding FGFa, SCF, IGF2, TPO, and heparin to the EB factors. After 10 days, hematopoietic progenitors were harvested and replated into fresh Stemline II medium supplemented with BMP4, SCF, Flt3 ligand, IL-3, IL-11, and erythropoietin (EPO) to direct differentiation along the erythroid lineage and to support extensive proliferation. After 17 days, cells were transferred into Stemline II medium containing a more specific erythroid cocktail that included insulin, transferrin, SCF, IGF1, IL-3, IL-11, and EPO for 7 days. In a final maturation step of 7 days (days 24–31), cells were transferred into IMDM with insulin, transferrin, and BSA and supplemented with EPO. Cells were harvested for analysis on days 10, 17, 24, and 31.

SUPPLEMENTAL INFORMATION

Supplemental information includes Supplemental Experimental procedures, seven figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2015.10.002.

AUTHOR CONTRIBUTIONS

Z.G., O.Q.-B., A.M.C., E.O., C.H., X. A., I.O., L.C., O.A., F.R.-F., E.G.-S., G.G., and J.C.S. performed experiments. Z.G., O.Q.-B, E.O., P.K., C.H., C.K., X.A., E.G.-S., F.P., S.J., J.M., B.R.D., and J.C.S. analyzed results. L.P., R.G., N.E, T.M.M., M.L.R., J.S., and A.G. designed reagents and provided samples. Z.G., O.Q.-B., J.A.B, B.R.D., and J.C.S. designed research. O.Q.-B. and J.C.S. wrote the paper. J.C.S. obtained financial support.

ACKNOWLEDGMENTS

The authors would like to thank Aurora de la Cal and Sergio Losada for their collaboration in the administrative work; Miguel Angel Martín, Jesús Martínez, and Edilia de Almeida for the careful maintenance of the animals; J. Bazinet for providing logistics support for the metabolomics experiments; and Norman Feltz for reading and correcting the manuscript. This work was supported by the following grants: Dirección General de Investigación de la Comunidad de Madrid (CellCAM; Ref S2010/BMD-2420), Fondo de Investigaciones Sanitarias, Instituto de Salud Carlos III (RETICS-RD12/0019/0023 and RD12/0019/0031), SAF2014-54885-R and SAF2012-31142 from MINECO, the Cystic Fibrosis Foundation (DAVIS12GO), and the National Heart Lung and Blood Institute (RC1HL099559). J.M. and E.O. are supported by the Scottish National Blood Transfusion Service and the Wellcome Trust Translation Award Grant 102610. Z.G. was supported by a fellowship of the Gobierno Vasco (Spain). R.G., L.P., and A.G. are employees of Cellectis. All relevant data are within the paper and its Supplemental Information files; to discuss access to materials, please contact the corresponding author.

Received: February 6, 2015
Revised: October 2, 2015
Accepted: October 5, 2015
Published: November 5, 2015

REFERENCES

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilić, J., Pekarik, V., Tiscornia, G., et al. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat. Biotechnol. *26*, 1276–1284.

Abyzov, A., Mariani, J., Palejev, D., Zhang, Y., Haney, M.S., Tomasini, L., Ferrandino, A.F., Rosenberg Belmaker, L.A., Szekely, A., Wilson, M., et al. (2012). Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature *492*, 438–442.

Amabile, G., Welner, R.S., Nombela-Arrieta, C., D'Alise, A.M., Di Ruscio, A., Ebralidze, A.K., Kraytsberg, Y., Ye, M., Kocher, O., Neuberg, D.S., et al. (2013). In vivo generation of transplantable human hematopoietic cells from induced pluripotent stem cells. Blood *121*, 1255–1264.

Asp, J., Persson, F., Kost-Alimova, M., and Stenman, G. (2006). CHCHD7-PLAG1 and TCEA1-PLAG1 gene fusions resulting from

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

cryptic, intrachromosomal 8q rearrangements in pleomorphic salivary gland adenomas. Genes Chromosomes Cancer *45*, 820–828.

Ban, H., Nishishita, N., Fusaki, N., Tabata, T., Saeki, K., Shikamura, M., Takada, N., Inoue, M., Hasegawa, M., Kawamata, S., and Nishikawa, S. (2011). Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. Proc. Natl. Acad. Sci. USA *108*, 14234–14239.

Beutler, E., and Gelbart, T. (2000). Estimating the prevalence of pyruvate kinase deficiency from the gene frequency in the general white population. Blood *95*, 3585–3588.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. Science *326*, 1509–1512.

Carroll, D. (2011). Genome engineering with zinc-finger nucleases. Genetics *188*, 773–782.

Cavazza, A., Moiani, A., and Mavilio, F. (2013). Mechanisms of retroviral integration and mutagenesis. Hum. Gene Ther. *24*, 119–131.

Wellcome Trust Sanger Institute (2014). COSMIC: catalog of somatic mutations in cancer. http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/.

De Gobbi, M., Viprakasit, V., Hughes, J.R., Fisher, C., Buckle, V.J., Ayyub, H., Gibbons, R.J., Vernimmen, D., Yoshinaga, Y., de Jong, P., et al. (2006). A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. Science *312*, 1215–1217.

Deyle, D.R., Li, L.B., Ren, G., and Russell, D.W. (2014). The effects of polymorphisms on human gene targeting. Nucleic Acids Res. *42*, 3119–3124.

Fermo, E., Bianchi, P., Chiarelli, L.R., Cotton, F., Vercellati, C., Writzl, K., Baker, K., Hann, I., Rodwell, R., Valentini, G., and Zanella, A. (2005). Red cell pyruvate kinase deficiency: 17 new mutations of the PK-LR gene. Br. J. Haematol. *129*, 839–846.

Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009). Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proc. Jpn. Acad., Ser. B, Phys. Biol. Sci. *85*, 348–362.

Garate, Z., Davis, B.R., Quintana-Bustamante, O., and Segovia, J.C. (2013). New frontier in regenerative medicine: site-specific gene correction in patient-specific induced pluripotent stem cells. Hum. Gene Ther. *24*, 571–583.

Genovese, P., Schiroli, G., Escobar, G., Di Tomaso, T., Firrito, C., Calabria, A., Moi, D., Mazzieri, R., Bonini, C., Holmes, M.C., et al. (2014). Targeted genome editing in human repopulating haematopoietic stem cells. Nature *510*, 235–240.

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., et al. (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature *471*, 63–67.

Hussein, S.M., Batada, N.N., Vuoristo, S., Ching, R.W., Autio, R., Närvä, E., Ng, S., Sourour, M., Hämäläinen, R., Olsson, C., et al. (2011). Copy number variation and selection during reprogramming to pluripotency. Nature *471*, 58–62.

Karakikes, I., Stillitano, F., Nonnenmacher, M., Tzimas, C., Sanoudou, D., Termglinchan, V., Kong, C.W., Rushing, S., Hansen, J., Ceholski, D., et al. (2015). Correction of human phospholamban R14del mutation associated with cardiomyopathy using targeted nucleases and combination therapy. Nat. Commun. *6*, 6955.

Loh, Y.H., Agarwal, S., Park, I.H., Urbach, A., Huo, H., Heffner, G.C., Kim, K., Miller, J.D., Ng, K., and Daley, G.Q. (2009). Generation of induced pluripotent stem cells from human blood. Blood *113*, 5476–5479.

Meza, N.W., Alonso-Ferrero, M.E., Navarro, S., Quintana-Bustamante, O., Valeri, A., Garcia-Gomez, M., Bueren, J.A., Bautista, J.M., and Segovia, J.C. (2009). Rescue of pyruvate kinase deficiency in mice by gene therapy using the human isoenzyme. Mol. Ther. *17*, 2000–2009.

Nishimura, K., Sano, M., Ohtaka, M., Furuta, B., Umemura, Y., Nakajima, Y., Ikehara, Y., Kobayashi, T., Segawa, H., Takayasu, S., et al. (2011). Development of defective and persistent Sendai virus vector: a unique gene delivery/expression system ideal for cell reprogramming. J. Biol. Chem. *286*, 4760–4771.

Nishishita, N., Takenaka, C., Fusaki, N., and Kawamata, S. (2011). Generation of human induced pluripotent stem cells from cord blood cells. J. Stem Cells *6*, 101–108.

Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. Nature *451*, 141–146.

Porteus, M.H., and Carroll, D. (2005). Gene targeting using zinc finger nucleases. Nat. Biotechnol. *23*, 967–973.

Renkawitz, J., Lademann, C.A., and Jentsch, S. (2014). Mechanisms and principles of homology search during recombination. Nat. Rev. Mol. Cell Biol. *15*, 369–383.

Rio, P., Baños, R., Lombardo, A., Quintana-Bustamante, O., Alvarez, L., Garate, Z., Genovese, P., Almarza, E., Valeri, A., Díez, B., et al. (2014). Targeted gene therapy and cell reprogramming in Fanconi anemia. EMBO Mol. Med. *6*, 835–848.

Sebastiano, V., Maeder, M.L., Angstman, J.F., Haddad, B., Khayter, C., Yeo, D.T., Goodwin, M.J., Hawkins, J.S., Ramirez, C.L., Batista, L.F., et al. (2011). In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. Stem Cells *29*, 1717–1726.

Seki, T., Yuasa, S., Oda, M., Egashira, T., Yae, K., Kusumoto, D., Nakata, H., Tohyama, S., Hashimoto, H., Kodaira, M., et al. (2010). Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. Cell Stem Cell *7*, 11–14.

Song, B., Fan, Y., He, W., Zhu, D., Niu, X., Wang, D., Ou, Z., Luo, M., and Sun, X. (2015). Improved hematopoietic differentiation efficiency of gene-corrected beta-thalassemia induced pluripotent stem cells by CRISPR/Cas9 system. Stem Cells Dev. *24*, 1053–1065.

Staerk, J., Dawlaty, M.M., Gao, Q., Maetzel, D., Hanna, J., Sommer, C.A., Mostoslavsky, G., and Jaenisch, R. (2010). Reprogramming of human peripheral blood cells to induced pluripotent stem cells. Cell Stem Cell *7*, 20–24.

Please cite this article in press as: Garate et al., Generation of a High Number of Healthy Erythroid Cells from Gene-Edited Pyruvate Kinase Deficiency Patient-Specific Induced Pluripotent Stem Cells, Stem Cell Reports (2015), http://dx.doi.org/10.1016/j.stemcr.2015.10.002

Suvatte, V., Tanphaichitr, V.S., Visuthisakchai, S., Mahasandana, C., Veerakul, G., Chongkolwatana, V., Chandanayingyong, D., and Issaragrisil, S. (1998). Bone marrow, peripheral blood and cord blood stem cell transplantation in children: ten years' experience at Siriraj Hospital. Int. J. Hematol. *68*, 411–419.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Tanphaichitr, V.S., Suvatte, V., Issaragrisil, S., Mahasandana, C., Veerakul, G., Chongkolwatana, V., Waiyawuth, W., and Ideguchi, H. (2000). Successful bone marrow transplantation in a child with red blood cell pyruvate kinase deficiency. Bone Marrow Transplant. *26*, 689–690.

Ye, Z., Zhan, H., Mali, P., Dowey, S., Williams, D.M., Jang, Y.Y., Dang, C.V., Spivak, J.L., Moliterno, A.R., and Cheng, L. (2009). Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. Blood *114*, 5473–5480.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917–1920.

Zanella, A., Fermo, E., Bianchi, P., and Valentini, G. (2005). Red cell pyruvate kinase deficiency: molecular and clinical aspects. Br. J. Haematol. *130*, 11–25.

Zanella, A., Bianchi, P., and Fermo, E. (2007). Pyruvate kinase deficiency. Haematologica *92*, 721–723.

Zou, J., Mali, P., Huang, X., Dowey, S.N., and Cheng, L. (2011). Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. Blood *118*, 4599–4608.


Hematopoietic
Innovative Therapies


VANCED
Therapies Unit


GOBIERNO
DE ESPAÑA
MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD
Ciemat
Centro de Investigaciones
Energéticas, Medioambientales
y Tecnológicas


er
ciberer
CENTRO DE INVESTIGACIÓN BIOMÉDICA EN RED
DE ENFERMEDADES RARAS


iis
FJD
INSTITUTO DE
INVESTIGACIÓN
SANITARIA
FUNDACIÓN JIMÉNEZ DÍAZ