

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2017322511 B2**

(54) Title
Gene therapy for patients with Fanconi anemia

(51) International Patent Classification(s)
C12N 15/85 (2006.01) **C12Q 1/68** (2006.01)
C12N 15/63 (2006.01)

(21) Application No: **2017322511** (22) Date of Filing: **2017.09.08**

(87) WIPO No: **WO18/049273**

(30) Priority Data

(31) Number	(32) Date	(33) Country
62/385,185	2016.09.08	US
62/412,028	2016.10.24	US

(43) Publication Date: **2018.03.15**

(44) Accepted Journal Date: **2021.08.26**

(71) Applicant(s)
Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas, O.A., M.P.;Fundacion Instituto de Investigacion Sanitaria Fundacion Jimenez Diaz;Consorcio Centro de Investigacion Biomedica en Red, M.P.;Fundacion para la Investigacion Biomedica del Hospital Infantil Universitario Nino Jesus

(72) Inventor(s)
Bueren, Juan A.;Rio, Paula;Navarro, Susana;Sevilla, Julian;Segovia, Jose Carlos;Gonzalez, Africa;Casado, Jose Antonio;Guenechea, Guillermo

(74) Agent / Attorney
Davies Collison Cave Pty Ltd, Level 15 1 Nicholson Street, MELBOURNE, VIC, 3000, AU

(56) Related Art
AFRICA GONZALEZ-MURILLO ET AL, "Development of Lentiviral Vectors with Optimized Transcriptional Activity for the Gene Therapy of Patients with Fanconi Anemia", HUMAN GENE THERAPY, (2010-05-01), vol. 21, no. 5, pages 623 - 630
JAKUB TOLAR ET AL, "Stem Cell Gene Therapy for Fanconi Anemia: Report from the 1st International Fanconi Anemia Gene Therapy Working Group Meeting", MOLECULAR THERAPY (2011-07-01), vol. 19, no. 7, pages 1193 - 1198
P S BECKER ET AL, "Preclinical correction of human Fanconi anemia complementation group A bone marrow cells using a safety-modified lentiviral vector", GENE THERAPY, GB, (2010-05-20), vol. 17, no. 10, pages 1244 - 1252



(51) International Patent Classification:

C12N 15/85 (2006.01) C12Q 1/68 (2006.01)
C12N 15/63 (2006.01)

(21) International Application Number:

PCT/US2017/050837

(22) International Filing Date:

08 September 2017 (08.09.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/385,185 08 September 2016 (08.09.2016) US
62/412,028 24 October 2016 (24.10.2016) US

(71) Applicants: CENTRO DE INVESTIGACIONES ENERGETICAS MEDIOAMBIENTALES Y TECNOLOGICAS [ES/ES]; Avda. de la Complutense, 40, 28040 Madrid (ES). FUNDACION INSTITUTO DE INVESTIGACION SANITARIA FUNDACION JIMENEZ DIAZ [ES/ES]; Avda. de los Reyes Catolicos, 2, 28040 Madrid (ES). CENTRO DE INVESTIGACION BIOMEDICA EN RED [ES/ES]; Monforte de Lemos, 5, 28029 Madrid (ES). FUNDACION PARA LA INVESTIGACION BIOMEDICA HOSPITAL INFAN-

TIL UNIVERSITARIO NINO JESUS [ES/ES]; Avda Menendez Pelayo 65, 28009 Madrid (ES).

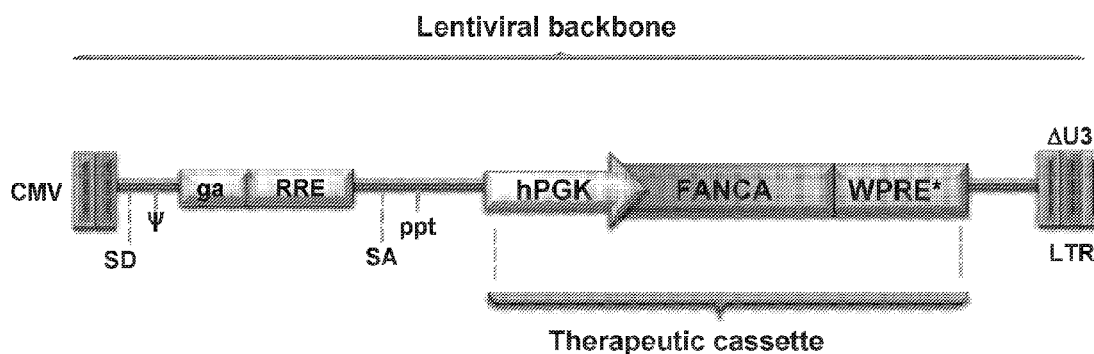
(71) Applicant (for OM only): SHERIN, Mridula [US/US]; 1700 Seventh Avenue, Suite 1900, Seattle, Washington 98101 (US).

(72) Inventors; and

(71) Applicants: BUEREN, Juan A. [ES/US]; c/o RTW Investments, LLC, 250 W 55th St., Suite A, New York, New York 10019 (US). RIO, Paula [ES/US]; c/o RTW Investments, LLC, 250 W 55th St., Suite A, New York, New York 10019 (US). NAVARRO, Susana [ES/US]; c/o RTW Investments, LLC, 250 W 55th St., Suite A, New York, New York 10019 (US). SEVILLA, Julian [ES/US]; c/o RTW Investments, LLC, 250 W 55th St., Suite A, New York, New York 10019 (US). SEGOVIA, Jose Carlos [ES/US]; c/o RTW Investments, LLC, 250 W 55th St., Suite A, New York, New York 10019 (US). GONZALEZ, Africa [ES/US]; c/o RTW Investments, LLC, 250 W 55th St., Suite A, New York, New York 10019 (US). CASADO, Jose Antonio [ES/US]; c/o RTW Investments, LLC, 250 W 55th St., Suite A, New York, New York 10019 (US). GUENECHEA, Guillermo [ES/US]; c/o RTW Investments, LLC, 250 W 55th St., Suite A, New York, New York 10019 (US).

(54) Title: GENE THERAPY FOR PATIENTS WITH FANCONI ANEMIA

Figure 1.



(57) Abstract: The present invention provides compositions and methods for rescuing FANCA expression in cells with diminished or no FANCA gene product. In particular, methods and compositions for gene therapy of Fanconi anemia are disclosed.

(74) **Agent:** SHERIN, Mridula et al.; Cooley LLP, 1299 Pennsylvania Avenue, NW, Suite 700, Washington, District of Columbia 20004 (US).

(81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

GENE THERAPY FOR PATIENTS WITH FANCONI ANEMIA

RELATED APPLICATIONS

[0001] This application claims priority to, and the benefit of, U.S. Provisional Application No. 62/385185, filed on September 8, 2016 and U.S. Provisional Application 62/412.028 filed October 24, 2016, the contents of which are incorporated herein by reference in their entireties.

FIELD OF THE INVENTION

[0002] The present invention relates generally to gene transfer into cells with diminished or no protein activity from one or more *FANCA* encoded proteins.

STATEMENT REGARDING SEQUENCE LISTING

[0003] The Sequence Listing associated with this application is provided in text format in lieu of a paper copy and is hereby incorporated by reference into the specification. The name of the text file containing the sequence is ROPA_002_01WO_ST25.txt. The text file is 46KB, was created on September 8, 2017, and is being submitted electronically via EFS-WEB.

BACKGROUND OF THE INVENTION

[0004] Fanconi Anemia (FA) is an autosomal recessive disease (except for complementation group FA-B, which is X-linked), where the median survival of patients is around 24 years (Butturini A, *et al.* (1994) Blood 84:1650-1655; Kutler DI, *et al.* (2003) Blood 101:1249-1256). At birth, the blood count of these patients is generally normal. Macrocytosis is often the first hematological abnormality detected in these patients. This usually evolves with thrombocytopenia, anemia and pancytopenia. Bone marrow failure (BMF) is usually observed in these patients after 5-10 years, with an average age of hematologic disease onset of 7 years. About 80% of patients with FA will develop evidence of BMF in the first decade of life. Based on epidemiological studies to date, if malignant episodes do not appear before aplasia, virtually all patients with FA will develop BMF by 40 years of age (Butturini A, *et al.* (1994) Blood 84:1650-1655; Kutler DI, *et al.* (2003) Blood 101:1249-1256), this being the leading cause of mortality in these patients.

[0005] Due to the complex clinical manifestations of FA, management of these patients is mainly focused on improving the following syndromes: bone marrow failure (BMF), myeloid leukemia, and solid tumors.

[0006] Current treatments include androgens such as those composed of fluoxymesterone, oxymetholone or stanozolol. As recently reviewed by Dufour and colleagues (Dufour, C. and Svahn, J. (2008). *Bone Marrow Transplant* 41 Suppl 2: S90-95), FA patients may have some response to androgens as long as this treatment is not initiated at a very advanced stage of the disease. About 75% of patients respond to androgens. A combination with 2 mg/kg/day of prednisolone reduces the risk of liver toxicity. In general, in the absence of a suitable bone marrow donor, androgens can be considered as an alternative treatment when there is some residual haematopoiesis, but not as a definitive long-term treatment.

[0007] Tischkowitz *et al.* noted that although in the early stages of the disease FA patients might respond to androgens, the vast majority of FA patients are refractory to these treatments in the long term (Tischkowitz, M. and Dokal, I. (2004). *Br J Haematol* 126: 176-191).

[0008] There are no specific indications for hematopoietic growth factors in the treatment of FA. However, two pharmacological groups of drugs with indications for specific symptoms of FA (anemia and neutropenia) have been identified: 1) erythropoietin and 2) granulocyte-colony stimulating factors (G-CSFs). Several erythropoietins (*e.g.*, Aranesp, Nespo, Exjade) are approved for the treatment of anemia, and G-CSFs (*e.g.*, G-CSF analogs such as filgrastim, biogastin, neulasta) are approved for the treatment of neutropenia. Although hematopoietic growth factors, such as erythropoietin and granulocyte colony-stimulating factors, have been tested in a limited number of patients with FA, responses were partial and transient (Dufour *et al.*, 2008). At present, these treatments do not represent good long-term options. For short-term treatment in neutropenic patients, G-CSF can be used for acute infections to increase the number of peripheral neutrophils, potentiating antibiotics. However, these drug treatments are far from definitive management for patients with FA.

[0009] Presently, the only curative treatment of the hematological manifestations of the disease is based on allogeneic hematopoietic transplantation. While the outcome of FA patients transplanted with grafts from HLA-identical sibling donors is in general satisfactory, only about 20% of FA patients will have an HLA-identical sibling. A significant proportion of FA patients without a sibling donor can be transplanted from

alternative donors, although these transplants are associated with a higher morbidity and mortality. In the remaining FA patients, no alternative therapies are currently available.

[00010] Accordingly, there remains a critical need for an effective treatment regimen for FA. The present invention addresses this need and more.

SUMMARY OF THE INVENTION

[00011] Embodiments of the present invention comprise polynucleotide cassettes for the enhanced expression of FANCA. In some embodiments, the polynucleotide cassette comprises a sequence encoding a codon-optimized human *FANCA* cDNA to increase mRNA stability upon transcription.

[00012] In one embodiment, the present invention includes an expression cassette comprising a polynucleotide sequence comprising in the following 5' to 3' order: (a) a human phosphoglycerate kinase (PGK) promoter sequence or a functional homolog or variant thereof; (b) a sequence encoding a human FANCA polypeptide or a functional fragment or variant thereof; (c) a woodchuck hepatitis virus regulatory element (WPRE) RNA export signal sequence or a functional variant or fragment thereof, wherein the sequence encoding the human FANCA polypeptide or functional fragment or variant thereof is operably linked to the PGK promoter sequence. In particular embodiments, the FANCA polypeptide or functional fragment or variant thereof comprises the sequence set forth in SEQ ID NO: 25; the sequence encoding the FANCA polypeptide or functional fragment or variant thereof comprises the sequence set forth in SEQ ID NO: 8; the PGK promoter comprises a nucleotide sequence of SEQ ID NO: 7; and/or the WPRE element comprises a nucleotide sequence of SEQ ID NO: 23. In particular embodiments, the cassette comprises a region of the nucleotide sequence of SEQ ID NO: 24. In certain embodiments, the cassette further comprises one or more enhancer sequences, a polypurine tract (PPT) or polyadenylation (poly A) signal sequence, a packing signal sequence, a truncated Gag sequence, a Rev responsive element (RRE), a central polypurine tract (cPPT), a central terminal sequence (CTS) and/or an upstream sequence element (USE), optionally from simian virus 40 (SV40-USE).

[00013] In one embodiment, the present invention provides an expression cassette comprising a polynucleotide sequence comprising: a) a promoter sequence; b) a sequence encoding a polypeptide; and c) a ribonucleic acid (RNA) export signal, wherein the promoter sequence is operably linked to the sequence encoding the FANCA polypeptide (SEQ ID NO: 25), and optionally where a)-c) are present in the expression cassette in 5' to

3' order. In certain embodiments, the promoter is a phosphoglycerate kinase (PGK) promoter. In certain embodiments, the sequence encoding the polypeptide is codon-optimized. In some embodiments, the sequence encoding the polypeptide is a codon-optimized version of the human FANCA cDNA having at least 85% identity to SEQ ID NO: 8. In particular embodiments, the RNA export signal is a mutated post-transcriptional regulatory element of the woodchuck hepatitis virus (wPRE).

[00014] In certain embodiments, the mutated wPRE is a chimeric wPRE comprising a sequence having at least 80% identity to SEQ ID NO: 23. In some embodiments, the expression cassette further comprising one or more enhancer sequences. In some embodiments, the expression cassette further comprises a polypurine tract (PPT) or polyadenylation (polyA) signal sequence. In some embodiments, the expression cassette further comprises one or more of the following sequences: i) a packing signal sequence; ii) a truncated Gag sequence; iii) a Rev responsive element (RRE); iv) a central polypurine tract (cPPT); v) a central terminal sequence (CTS); and vi) an upstream sequence element (USE), optionally from simian virus 40 (SV40-USE). In some embodiments, the expression cassette further comprises 5' and 3' long terminal repeat (LTR) sequences.

[00015] In a related embodiment, the present invention provides a recombinant gene delivery vector comprising an expression cassette disclosed herein. In certain embodiments, the recombinant gene delivery vector is a virus or viral vector. In certain embodiments, the virus or viral vector is a lentivirus (LV).

[00016] In another related embodiment, the present invention provides a cell comprising an expression cassette or gene delivery vector disclosed herein. In some embodiments, the cell is a blood cell. In some embodiments, the cell is an erythroid cell. In some embodiments, the cell is a bone marrow cell, e.g., a lineage depleted bone marrow cell. In particular embodiments, the cell is a hematopoietic stem cell or a CD34⁺ cell. In some embodiments, the cell is a hematopoietic stem cell. In some embodiments, the cell is a CD34⁺ hematopoietic stem cell. In some embodiments, the cell is a committed hematopoietic erythroid progenitor cell.

[00017] In a related embodiment, the present invention provides a pharmaceutical composition comprising a pharmaceutically acceptable excipient and recombinant gene delivery vector or cell disclosed herein. In certain aspects of the invention, pharmaceutical compositions are provided comprising a polynucleotide cassette of the invention and a

pharmaceutical excipient. In other embodiments, the pharmaceutical composition comprises a gene delivery vector of the invention and a pharmaceutical excipient.

[00018] Methods and compositions are provided for the use of gene therapy vector compositions, *e.g.*, viral vectors, comprising these genetic expression cassettes for use in the preparation of medicaments useful in central and targeted gene therapy of diseases, disorders, and dysfunctions in an animal, and in humans in particular.

[00019] In another embodiment, the present invention provides a method of treating or preventing a disease or disorder in a subject in need thereof, comprising providing to the subject an expression cassette, gene delivery vector, or pharmaceutical composition disclosed herein.

[00020] In another embodiment, the present invention includes a method of treating Fanconi anemia in a subject in need thereof, comprising providing to the subject a pharmaceutical composition disclosed herein.

[00021] In a related embodiment, the present invention includes a method for treating Fanconi anemia in a subject in need thereof, comprising providing to the subject CD34⁺ cells comprising an expression cassette, wherein the expression cassette comprises a polynucleotide sequence comprising in the following 5' to 3' order: (a) a human phosphoglycerate kinase (PGK) promoter sequence or a functional homolog or variant thereof; (b) a sequence encoding a human FANCA polypeptide or a functional fragment or variant thereof; (c) a woodchuck hepatitis virus regulatory element (WPRE) RNA export signal sequence or a functional variant or fragment thereof, wherein the sequence encoding the human FANCA polypeptide or functional fragment or variant thereof is operably linked to the PGK promoter sequence. In certain embodiments, the CD34⁺ cells were obtained from the subject. In particular embodiments, the CD34⁺ cells were obtained from the subject after the subject was treated with a combination of: (i) G-CSF or Filgrastin; and (ii) Plerifaxor. In particular embodiments, the CD34⁺ cells were transduced with the recombinant gene delivery vector comprising the expression cassette. In one embodiment, the CD34⁺ cells were transduced by contacting the CD34⁺ cells with the recombinant gene delivery vector for about 24 hours.

[00022] In another embodiment, the present invention provides a method for treating Fanconi anemia in a subject in need thereof, comprising: (a) providing to the subject a combination of: (i) G-CSF or Filgrastin; and (ii) Plerifaxor to mobilize CD34⁺ cells within the subject; (b) obtaining a biological sample comprising CD34⁺ cells from the subject, wherein the biological sample is

optionally peripheral blood or bone marrow; (c) preparing a cell population enriched for CD34⁺ cells from the biological sample; (d) transducing the cell population enriched for CD34⁺ cells with a recombinant gene deliver vector comprising an expression cassette comprising a polynucleotide sequence comprising in the following 5' to 3' order: (i) a promoter sequence or a functional homolog or variant thereof; and (ii) a sequence encoding a human FANCA polypeptide or a functional fragment or variant thereof, wherein the sequence encoding the human FANCA polypeptide or functional fragment or variant thereof is operably linked to the PGK promoter sequence, where the transducing comprises contacting the cell population enriched for CD34⁺ cells with the lentiviral vector for about 24 hours; and (e) providing the cell population transduced with the lentiviral vector resulting from step (d) to the subject. In certain embodiments, preparing the cell population comprises depleting erythrocytes and/or enriching for CD34⁺ cells by positive selection, negative selection, or a combination thereof. In particular embodiments, the method inhibits the development of, halts progression of, and/or reverses progression of a hematological manifestation of Fanconi anemia in the subject. In particular embodiments, the hematological manifestation of Fanconi anemia is selected from one or more of BMF, thrombocytopenia, leukopenia, pancytopenia, neutropenia, and anemia.

[00023] Other features and advantages of the invention will be apparent from and encompassed by the following detailed description and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[00024] **Figure 1** is schematic diagram of an exemplary construct, PGK-FANCA.WPRE*LV.

[00025] **Figure 2A** shows a schematic representation of LVs expressing FANCA under the control of different internal promoters. **Figure 2B** shows a Western blot analysis of FANCA in FA-A cells transduced with vectors shown in panel A. shows correction of the MMC-hypersensitivity of hematopoietic progenitors from FA-A mice subjected to gene therapy. FA-A bone marrow (BM) cells were transduced with PGK_FANCA-WPRE* or control SF1-EGFP LVs and transplanted into irradiated FA-A mice. At 7 months post-transplantation BM samples were harvested and cultured in methylcellulose in the presence of increasing concentrations of mitomycin C (MMC).

[00026] **Figure 3** presents data showing the functional analysis of lentiviral vectors expressing FANCA under the control of different internal promoters. **Figure 3A** shows

reversion of MMC sensitivity of FA-A lymphoblast cell line (LCL) cells transduced with LVs. Mean values of 3 different experiments are shown. **Figure 3B** show restored formation of nuclear FANCD2 foci in FA-A LCLs transduced with vectors and exposed to mitomycin C (MMC).

[00027] **Figure 4** presents data showing *in vivo* efficacy and safety of FA gene therapy with the PGK-FANCA.Wpre* LV **Figure 4A** shows the construct. **Figure 4B** depicts the methodology whereby Bone marrow (BM) cells from FA-A mice were transduced with FANCA LV and then transplanted into irradiated FA-A recipient mice. **Figure 4C** shows BM samples from transplanted FA-A mice were cultured in methylcellulose in the absence and the presence of MMC.

[00028] **Figure 5** shows proviral copy number in FANCA^{-/-} mice transplanted with syngenic bone marrow cells previously transduced with lentiviral vectors carrying the therapeutic FANCA gene under the control of the PGK promoter. PB = peripheral blood; BM = bone marrow.

[00029] **Figure 6** presents data showing correction of the MMC-hypersensitivity of hematopoietic progenitors from FAA mice subjected to gene therapy with the medicinal product. FA-A bone marrow (BM) cells were transduced with PGK_FANCA-Wpre* or SF1-EGFP LVs and transplanted into irradiated FA-A mice. At seven (7) months post-transplantation BM samples were harvested and cultured in methylcellulose in the presence of increasing concentrations of MMC.

[00030] **Figure 7** depicts improved transduction efficacy of cryopreserved bone marrow progenitors from three Fanconi Anemia patients. Samples were subjected to standard transductions consisting in a single transduction cycle (16h) after 2h of static preloading (white bars; 1xS) or improved transduction consisting in three transduction cycles (2h+2h+12h) with the lentiviral vectors (grey bars; 3xD).

[00031] **Figure 8** shows the relevance of the WPRE sequence on the functional properties of lentiviral vectors expressing FANCA under the control of the PGK promoter.

Figure 8A: Reversion of the (MMC) sensitivity of FA-A LCLs transduced with PGK-FANCA and PGK-FANCA-WPRE LVs. Mean values of 3 different experiments are shown.

Figure 8B: Reversion of MMC sensitivity of FA-A hematopoietic progenitors (colony forming cells, CFCs) transduced with SFFV-FANCA LV and PGK-FANCA LVs ("Expt 1")

and with PGK-FANCA and PGK-FANCA-WPRE LVs (“Expt 2”). White bars = no MMC; Black bars = 10 nM MMC. MMC = mitomycin C

[00032] **Figure 9** shows efficacy of GALV-TR and VSV-G pseudo typed lentiviral vectors to transduce hematopoietic progenitors from the bone marrow of Fanconi anemia patients with EGFP-LVs.

[00033] **Figure 10** shows low *in vitro* transformation potential of lentiviral vectors harboring the hPGK promoter. **Figure 10A**: depiction of the vectors. **Figure 10B**: transformation capacity as measured in re-plating frequency over copy number.

[00034] **Figure 11** is a depiction of an illustrative hematopoietic stem cell (HSC) collection and gene therapy trial of FA-A patients.

[00035] **Figure 12** shows the hematological parameters of recruited patients in the study described in the Examples. **Figure 12A** shows results for hemoglobin. **Figure 12B** shows results for neutrophils. **Figure 12C** shows results for platelets. **Figure 12D** shows results for CD34+ cells.

[00036] **Figure 13** illustrates the Fancostem protocol phase II study aiming at the evaluation of the safety and efficacy of the mobilization and collection of CD34+ cells after treatment with Plerixafor (MOZOBIL) and Filgrastim (also known as G-CSF) (NEUPOGENE) in patients with Fanconi anemia. The number of patients is 10.

[00037] **Figure 14** shows G-CSF/Plerixafor-mediated mobilization of CD34+ cells in FA-A patients.

[00038] **Figure 15** shows G-CSF/Plerixafor-mediated mobilization of CFCs in FA-A patients.

[00039] **Figure 16** is a summary of the CD34+ cells collected in G-CSF/Plerixafor mobilized FA-A patients. **Figure 16A** shows CD34+ cell collection in FANCOSTEM and **Figure 16B** shows compared to previous studies.

[00040] **Figure 17** is a chart showing the comparison between predicted CD34+ cell numbers in bone marrow (BM) versus actual numbers in mobilized peripheral blood (mPB).

[00041] **Figure 18** is a chart of the collection and purification of mobilized peripheral blood (mPB) FA-A CD34+ cells.

[00042] **Figure 19** shows CD34 expression prior to and after immunoselection of mobilized peripheral blood (mPB) CD34+ cells from healthy donors (HD) and FA patients.

[00043] **Figure 20** shows patient FA 02005 fit the criteria for both FANCOSTEM and FANCOLEN studies. **Figure 20A** shows cell counts; **Figure 20B** shows hematopoietic stem cell (HSC) content versus age.

[00044] **Figure 21 (A-E)** present test results showing FA diagnosis of patient FA-02005 prior to gene therapy.

[00045] **Figure 22** shows the follow up parameter of the cell manufacturing process for FA-A Patient 02005.

[00046] **Figure 23** is a graph depicting vector copy number prior to and at 2 weeks, 4 weeks, 6 weeks, 2 months, 3 months, 4 months, and 5 months after gene therapy in patient FA-02005.

[00047] **Figure 24** presents follow-up of the first not-conditioned FA-A patient (FA-02005) prior to and after gene therapy as measured by hemoglobin amounts.

[00048] **Figure 25** presents follow-up of the first not-conditioned FA-A patient (FA-02005) prior to and after gene therapy as measured by neutrophil amounts.

[00049] **Figure 26** presents follow-up of the first not-conditioned FA-A patient(FA-02005) prior to and after gene therapy as measured by platelet amounts.

[00050] **Figure 27** is a chart of the hematological evolution of patient FA-A 02002.

[00051] **Figure 28** shows the diagnosis of FA 02002 as Not mosaic; Homozygote *FANCA* c.239 C>T p.Gln99*MMC Hypersensitive; Complemented by *FANC*.

[00052] **Figure 29** shows the cell manufacturing process in patient FA-A 02002.

[00053] **Figure 30** presents the analysis of CD34 expression in a healthy donor (HD) and FA mobilized peripheral blood (mPB) during the different steps required for LV-transduction in patient FA 02002.

[00054] **Figure 31** is a graph depicting vector copy number prior to and after gene therapy in patient FA 02002.

[00055] **Figure 32** presents data of the follow up of patient FA-A 02002 infused with cryopreserved cells as measured by hemoglobin amount.

[00056] **Figure 33** presents data of the follow up of patient FA-A 02002 infused with cryopreserved cells as measured by neutrophil amount.

[00057] **Figure 34** presents data of the follow up of patient FA-A 02005 infused with cryopreserved cells as measured by platelet amount.

[00058] **Figure 35** shows transduction of fresh mobilized peripheral blood (mPB) CD34+ cells from FA-A patients using validated conditions. **Figure 35A** presents the protocol.

Figure 35B shows a graph of results from patient 02002. **Figure 35C** shows a graph of results from patient 02003. **Figure 35D** shows a graph of results from patient 02004.

[00059] **Figure 36** shows data for the engraftment of corrected FA-A mPB CD34+ cells in NSG mice. **Figure 36A** shows the protocol. **Figure 36B** shows results for patient 02002. **Panel C** shows results for patient 02003. **Figure 36D** shows results for patient 02004. mPB = mobilized peripheral blood.

[00060] **Figure 37** depicts *in vivo* selection of corrected FA HPCs from patient 02002 in NOD scid gamma (NSG) mice. **Figure 37A** shows the protocol. **Figure 37B** is a graph of CFCs pre-transplantation. **Figure 37C** is a graph of hCFCs 30 days post transplantation.

[00061] **Figure 38** is a map of the 5.7kb plasmid encoding the envelope G glycoprotein of the VSV under the control of the CMV promoter and carries the kanamycin resistant gene for selection purposes.

[00062] **Figure 39** is a map of the 3.5kb plasmid encoding for the HIV-1 *rev* gene under the control of the CMV promoter and carries the kanamycin resistance gene for selection purposes.

[00063] **Figure 40** is a map of the 8.8kb plasmid containing the HIV-1 *gag* and *pol* genes that code for the HIV-1 structural and enzymatic proteins under the control of the CMV promoter. It contains intron 2 of the human beta globin (HBB2), the HIV-1 Rev responsive element (RRE) and the kanamycin resistance gene.

[00064] **Figure 41** is a map of the 11621 base pair transfer cassette pCCL-SIN-cPPT/CTS-hPGK-hFANCA-WPRE.

[00065] **Figure 42** represents the LAM-PCR analysis of FANCA-LV insertion sites in FA hematopoietic stem cells (HSC).

[00066] **Figure 43** depicts LAM-PCR results for tracking of FANCA-LV treated cells. **Figure 43A** depicts the protocol and **Figure 43B** is a chart of the data.

[00067] **Figure 44** shows the clonal diversity of Fanca -/- recipients transplanted with LV-corrected HSCs. **Figure 4A** shows the protocol and **Figure 44B** presents a graph of the data.

DETAILED DESCRIPTION OF THE INVENTION

[00068] The present invention relates generally to the fields of molecular biology and virology, and in particular, to gene expression cassettes, and vectors comprising them useful for the delivery of nucleic acid segments encoding selected therapeutic constructs (including for example, peptides, polypeptides, ribozymes, and catalytic RNA molecules), to selected cells and tissues of vertebrate animals. In particular, these genetic constructs are useful in gene therapy for the treatment of mammalian, and in particular, human diseases, disorders, and dysfunctions related to FANCA gene product dysregulation.

[00069] In certain embodiments, the invention provides compositions and methods for gene therapy treatment of subjects with Fanconi Anemia (FA). In particular, compositions and methods for rescuing FANCA gene expression are provided. Specific methods disclosed herein relate to the use of lentiviral vectors to deliver human FANCA to hematopoietic progenitor cells of a subject with FA, particularly FA-A.

DEFINITIONS

[00070] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although methods and materials similar or equivalent to those described herein can be used in the practice of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are expressly incorporated by reference in their entirety. In cases of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples described herein are illustrative only and are not intended to be limiting.

[00071] A "vector" as used herein refers to a macromolecule or association of macromolecules that comprises or associates with a polynucleotide and which can be used to mediate delivery of the polynucleotide to a cell. Illustrative vectors include, for example, plasmids, viral vectors (*e.g.*, retroviral vectors, such as lentiviral vectors), liposomes, and other gene delivery vehicles.

[00072] The term "LV" is an abbreviation for lentivirus, and may be used to refer to the virus itself or derivatives thereof. The term covers all subtypes and both naturally occurring and recombinant forms, except where required otherwise.

[00073] As used herein, the term "gene" or "coding sequence" refers to a nucleotide sequence *in vitro* or *in vivo* that encodes a gene product. In some instances, the gene

consists or consists essentially of coding sequence, that is, sequence that encodes the gene product. In other instances, the gene comprises additional, non-coding, sequence. For example, the gene may or may not include regions preceding and following the coding region, *e.g.*, 5' untranslated (5' UTR) or "leader" sequences and 3' UTR or "trailer" sequences, as well as intervening sequences (introns) between individual coding segments (exons).

[00074] As used herein, a "therapeutic gene" refers to a gene that, when expressed, confers a beneficial effect on the cell or tissue in which it is present, or on a mammal in which the gene is expressed. Examples of beneficial effects include amelioration of a sign or symptom of a condition or disease, prevention or inhibition of a condition or disease, or conferral of a desired characteristic. Therapeutic genes include genes that correct a genetic deficiency in a cell or mammal.

[00075] As used herein, a transgene is a gene that is delivered to a cell by a vector.

[00076] As used herein, the term "gene product" refers to the desired expression product of a polynucleotide sequence such as a polypeptide, peptide, protein or interfering RNA including short interfering RNA (siRNA), miRNA or small hairpin RNA (shRNA).

[00077] As used herein, the terms "polypeptide," "peptide," and "protein" refer to polymers of amino acids of any length. The terms also encompass an amino acid polymer that has been modified; for example, disulfide bond formation, glycosylation, lipidation, phosphorylation, or conjugation with a labeling component.

[00078] By "comprising" it is meant that the recited elements are required in, for example, the composition, method, kit, etc., but other elements may be included to form the, for example, composition, method, kit etc. within the scope of the claim. For example, an expression cassette "comprising" a gene encoding a therapeutic polypeptide operably linked to a promoter is an expression cassette that may include other elements in addition to the gene and promoter, *e.g.*, poly-adenylation sequence, enhancer elements, other genes, linker domains, etc.

[00079] By "consisting essentially of", it is meant a limitation of the scope of the, for example, composition, method, kit, etc., described to the specified materials or steps that do not materially affect the basic and novel characteristic(s) of the, for example, composition, method, kit, etc. For example, an expression cassette "consisting essentially of" a gene encoding a therapeutic polypeptide operably linked to a promoter and a polyadenylation

sequence may include additional sequences, *e.g.*, linker sequences, so long as they do not materially affect the transcription or translation of the gene. As another example, a variant, or mutant, polypeptide fragment “consisting essentially of” a recited sequence has the amino acid sequence of the recited sequence plus or minus about 10 amino acid residues at the boundaries of the sequence based upon the full length naïve polypeptide from which it was derived, *e.g.*, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 residue less than the recited bounding amino acid residue, or 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 residues more than the recited bounding amino acid residue.

[00080] By “consisting of”, it is meant the exclusion from the composition, method, or kit of any element, step, or ingredient not specified in the claim. For example, an expression cassette “consisting of” a gene encoding a therapeutic polypeptide operably linked to a promoter, and a post-transcriptional regulatory element consists only of the promoter, polynucleotide sequence encoding the therapeutic polypeptide, and post-transcriptional regulatory element. As another example, a polypeptide “consisting of” a recited sequence contains only the recited sequence.

[00081] An “expression vector” as used herein encompasses a vector, *e.g.*, plasmid, mini-circle, viral vector, liposome, and the like as discussed above or as known in the art, comprising a polynucleotide which encodes a gene product of interest, and is used for effecting the expression of a gene product in an intended target cell. An expression vector also comprises control elements operatively linked to the encoding region to facilitate expression of the gene product in the target. The combination of control elements, *e.g.*, promoters, enhancers, UTRs, miRNA targeting sequences, etc., and a gene or genes to which they are operably linked for expression is sometimes referred to as an “expression cassette.” Many such control elements are known and available in the art or can be readily constructed from components that are available in the art.

[00082] A “promoter” as used herein encompasses a DNA sequence that directs the binding of RNA polymerase and thereby promotes RNA synthesis, *i.e.*, a minimal sequence sufficient to direct transcription. Promoters and corresponding protein or polypeptide expression may be ubiquitous, meaning strongly active in a wide range of cells, tissues and species or cell-type specific, tissue-specific, or species specific. Promoters may be “constitutive,” meaning continually active, or “inducible,” meaning the promoter can be activated or deactivated by the presence or absence of biotic or abiotic factors. Also

included in the nucleic acid constructs or vectors of the invention are enhancer sequences that may or may not be contiguous with the promoter sequence. Enhancer sequences influence promoter-dependent gene expression and may be located in the 5' or 3' regions of the native gene.

[00083] An "enhancer" as used herein encompasses a cis-acting element that stimulates or inhibits transcription of adjacent genes. An enhancer that inhibits transcription also is termed a "silencer". Enhancers can function (*i.e.*, can be associated with a coding sequence) in either orientation, over distances of up to several kilobase pairs (kb) from the coding sequence and from a position downstream of a transcribed region.

[00084] A "termination signal sequence" as used herein encompasses any genetic element that causes RNA polymerase to terminate transcription, such as for example a polyadenylation signal sequence.

[00085] As used herein, the terms "operatively linked" or "operably linked" refers to a juxtaposition of genetic elements, *e.g.*, promoter, enhancer, termination signal sequence, polyadenylation sequence, etc., wherein the elements are in a relationship permitting them to operate in the expected manner. For instance, a promoter is operatively linked to a coding region if the promoter helps initiate transcription of the coding sequence. There may be intervening residues between the promoter and coding region so long as this functional relationship is maintained.

[00086] As used herein, the term "heterologous" means derived from a genotypically distinct entity from that of the rest of the entity to which it is being compared. For example, a polynucleotide introduced by genetic engineering techniques into a plasmid or vector derived from a different species is a heterologous polynucleotide. As another example, a promoter removed from its native coding sequence and operatively linked to a coding sequence with which it is not naturally found linked is a heterologous promoter. Thus, for example, an LV vector that includes a heterologous nucleic acid encoding a heterologous gene product is an LV vector that includes a nucleic acid not normally included in a naturally-occurring, wild-type LV, and the encoded heterologous gene product is a gene product not normally encoded by a naturally-occurring, wild-type LV.

[00087] The term "endogenous" as used herein with reference to a nucleotide molecule or gene product refers to a nucleic acid sequence, *e.g.*, gene or genetic element, or gene

product, *e.g.*, RNA, protein, that is naturally occurring in or associated with a host virus or cell.

[00088] The term "native" as used herein refers to a nucleotide sequence, *e.g.*, gene, or gene product, *e.g.*, RNA, protein, that is present in a wildtype virus or cell.

[00089] The term "variant" as used herein refers to a mutant of a reference polynucleotide or polypeptide sequence, for example a native polynucleotide or polypeptide sequence, *i.e.*, having less than 100% sequence identity with the reference polynucleotide or polypeptide sequence. Put another way, a variant comprises at least one amino acid difference (*e.g.*, amino acid substitution, amino acid insertion, amino acid deletion) relative to a reference polynucleotide sequence, *e.g.*, a native polynucleotide or polypeptide sequence. For example, a variant may be a polynucleotide having a sequence identity of 70% or more with a full length native polynucleotide sequence, *e.g.*, an identity of 75% or 80% or more, such as 85%, 90%, or 95% or more, for example, 98% or 99% identity with the full length native polynucleotide sequence. As another example, a variant may be a polypeptide having a sequence identity of 70% or more with a full length native polypeptide sequence, *e.g.*, an identity of 75% or 80% or more, such as 85%, 90%, or 95% or more, for example, 98% or 99% identity with the full length native polypeptide sequence. Variants may also include variant fragments of a reference, *e.g.*, native, sequence sharing a sequence identity of 70% or more with a fragment of the reference, *e.g.*, native, sequence, *e.g.*, an identity of 75% or 80% or more, such as 85%, 90%, or 95% or more, for example, 98% or 99% identity with the native sequence.

[00090] As used herein, the terms "biological activity" and "biologically active" refer to the activity attributed to a particular biological element in a cell. For example, the "biological activity" of an "immunoglobulin", "antibody" or fragment or variant thereof refers to the ability to bind an antigenic determinant and thereby facilitate immunological function. As another example, the biological activity of a polypeptide or functional fragment or variant thereof refers to the ability of the polypeptide or functional fragment or variant thereof to carry out its native functions of, *e.g.*, binding, enzymatic activity, etc. As a third example, the biological activity of a gene regulatory element, *e.g.*, promoter, enhancer, kozak sequence, and the like, refers to the ability of the regulatory element or functional fragment or variant thereof to regulate, *i.e.*, promote, enhance, or activate the translation of, respectively, the expression of the gene to which it is operably linked.

[00091] The terms "administering" or "introducing", as used herein, refer to delivery of a vector for recombinant protein expression to a cell, to cells and/or organs of a subject, or to a subject. Such administering or introducing may take place *in vivo*, *in vitro* or *ex vivo*. A vector for expression of a gene product may be introduced into a cell by transfection, which typically means insertion of heterologous DNA into a cell by physical means (*e.g.*, calcium phosphate transfection, electroporation, microinjection or lipofection); infection, which typically refers to introduction by way of an infectious agent, *i.e.*, a virus; or transduction, which typically means stable infection of a cell with a virus or the transfer of genetic material from one microorganism to another by way of a viral agent (*e.g.*, a bacteriophage).

[00092] "Transformation" is typically used to refer to bacteria comprising heterologous DNA or cells which express an oncogene and have therefore been converted into a continuous growth mode such as tumor cells. A vector used to "transform" a cell may be a plasmid, virus or other vehicle.

[00093] Typically, a cell is referred to as "transduced", "infected"; "transfected" or "transformed" dependent on the means used for administration, introduction or insertion of heterologous DNA (*i.e.*, the vector) into the cell. The terms "transduced", "transfected" and "transformed" may be used interchangeably herein regardless of the method of introduction of heterologous DNA.

[00094] The term "host cell", as used herein refers to a cell which has been transduced, infected, transfected or transformed with a vector. The vector may be a plasmid, a viral particle, a phage, etc. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to those skilled in the art. It will be appreciated that the term "host cell" refers to the original transduced, infected, transfected or transformed cell and progeny thereof.

[00095] The terms "treatment", "treating" and the like are used herein to generally mean obtaining a desired pharmacologic and/or physiologic effect. The effect may be prophylactic in terms of completely or partially preventing a disease or symptom thereof, *e.g.*, reducing the likelihood that the disease or symptom thereof occurs in the subject, and/or may be therapeutic in terms of a partial or complete cure for a disease and/or adverse effect attributable to the disease. "Treatment" as used herein covers any treatment of a disease in a mammal, and includes: (a) preventing the disease from occurring in a subject which may be predisposed to the disease but has not yet been diagnosed as having it; (b)

inhibiting the disease, *i.e.*, arresting its development; or (c) relieving the disease, *i.e.*, causing regression of the disease. The therapeutic agent may be administered before, during or after the onset of disease or injury. The treatment of ongoing disease, where the treatment stabilizes or reduces the undesirable clinical symptoms of the patient, is of particular interest. Such treatment is desirably performed prior to complete loss of function in the affected tissues. The subject therapy will desirably be administered during the symptomatic stage of the disease, and in some cases after the symptomatic stage of the disease.

[00096] The terms "individual," "host," "subject," and "patient" are used interchangeably herein, and refer to a mammal, including, but not limited to, human and non-human primates, including simians and humans; mammalian sport animals (*e.g.*, horses); mammalian farm animals (*e.g.*, sheep, goats, etc.); mammalian pets (dogs, cats, etc.); and rodents (*e.g.*, mice, rats, etc.). The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the invention. As used herein, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. Furthermore, to the extent that the terms "including", "includes", "having", "has", "with", or variants thereof are used in either the detailed description and/or the claims, such terms are intended to be inclusive in a manner similar to the term "comprising".

[00097] The term "about" or "approximately" means within an acceptable error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined, *i.e.*, the limitations of the measurement system. For example, "about" can mean within 1 or more than 1 standard deviation, per the practice in the art. Alternatively, "about" can mean a range of up to 20%, preferably up to 10%, more preferably up to 5%, and more preferably still up to 1% of a given value. Alternatively, particularly with respect to biological systems or processes, the term can mean within an order of magnitude, preferably within 5-fold, and more preferably within 2-fold, of a value. Where particular values are described in the application and claims, unless otherwise stated the term "about" meaning within an acceptable error range for the particular value should be assumed.

[00098] The practice of the present invention employs, unless otherwise indicated, conventional techniques of cell biology, molecular biology (including recombinant

techniques), microbiology, biochemistry and immunology, which are within the scope of those of skill in the art. Such techniques are explained fully in the literature, such as, "Molecular Cloning: A Laboratory Manual", second edition (Sambrook *et al.*, 1989); "Oligonucleotide Synthesis" (M. J. Gait, ed., 1984); "Animal Cell Culture" (R. I. Freshney, ed., 1987); "Methods in Enzymology" (Academic Press, Inc.); "Handbook of Experimental Immunology" (D. M. Weir & C. C. Blackwell, eds.); "Gene Transfer Vectors for Mammalian Cells" (J. M. Miller & M. P. Calos, eds., 1987); "Current Protocols in Molecular Biology" (F. M. Ausubel *et al.*, eds., 1987); "PCR: The Polymerase Chain Reaction", (Mullis *et al.*, eds., 1994); and "Current Protocols in Immunology" (J. E. Coligan *et al.*, eds., 1991), each of which is expressly incorporated by reference herein.

[00099] Several aspects of the invention are described below with reference to example applications for illustration. It should be understood that numerous specific details, relationships, and methods are set forth to provide a full understanding of the invention. One having ordinary skill in the relevant art, however, readily recognizes that the invention can be practiced without one or more of the specific details or with other methods. The present invention is not limited by the illustrated ordering of acts or events, as some acts may occur in different orders and/or concurrently with other acts or events. Furthermore, not all illustrated acts or events are required to implement a methodology in accordance with the present invention.

[000100] Unless otherwise indicated, all terms used herein have the same meaning as they would to one skilled in the art and the practice of the present invention will employ conventional techniques of microbiology and recombinant DNA technology, which are within the knowledge of those of skill of the art.

[000101] In certain embodiments, the present disclosure provides polynucleotides, polynucleotide cassettes and expression vectors for the expression of a gene in cells. Also provided are pharmaceutical compositions and methods for the use of any of the compositions in promoting the expression of a gene in cells, for example, in an individual, e.g. for the treatment or prophylaxis of a disorder. These and other objects, advantages, and features of the invention will become apparent to those persons skilled in the art upon reading the details of the compositions and methods as more fully described below.

[000102] In certain embodiments, methods and compositions are provided for preparation of gene therapy vector compositions, e.g., viral vectors, comprising these genetic expression

cassettes for use in the preparation of medicaments useful in central and targeted gene therapy of diseases, disorders, and dysfunctions in an animal, and in humans in particular.

[000103] In some embodiments, the present invention provides for gene therapy for Fanconi Anemia based on a LV vector harbouring the hPGK eukaryotic promoter that drives the expression of the *FANCA* cDNA. This therapeutic vector may be used to transduce human hematopoietic stem cells (HSCs), which may be subsequently transplanted into humans with Fanconi Anemia.

[000104] In certain embodiments, the present invention provides an FANCA LV vector for the genetic correction of Fanconi Anemia. Overall, results demonstrate feasibility of gene therapy for FA with a LV designed for clinical application.

[000105] The disclosed compositions may be utilized in a variety of investigative, diagnostic and therapeutic regimens, including the prevention and treatment of a variety of human diseases. The various compositions and methods of the invention are described below.

[000106] Although particular compositions and methods are exemplified herein, it is understood that any of a number of alternative compositions and methods are applicable and suitable for use in practicing the invention. It will also be understood that an evaluation of the expression constructs and methods of the invention may be carried out using procedures standard in the art.

FANCONI ANEMIA

[000107] Fanconi Anemia (FA) is a rare inherited chromosomal instability syndrome mainly characterized by bone marrow failure (BMF) and cancer predisposition (Butturini A *et al.* Blood. 1994;84:1650-1655; Kutler DI *et al.* Blood. 2003; 101:1249-1256.). The prevalence of FA is 1-5 per million, and the heterozygote carrier frequency is estimated to be 1 in 300 (Tamary H *et al.* Eur J Haematol. 2004;72:330-335).

[000108] FA is both genetically and phenotypically heterogeneous. To date, thirteen (13) complementation groups have been reported (FA-A, B, C, D1, D2, E, F, G, I, J, L, M and N) associated with mutations in the corresponding 13 Fanconia Anemia Complementation group (FANC) genes: *FANCA*, *FANCB*, *FANCC*, *FANCD1/BRCA2*, *FANCD2*, *FANCE*, *FANCF*, *FANCG/XRCC9*, *FANCI*, *BRIP1/FANCI*, *FANCL*, *FANCM/Hef* and *FANCN/PALB2* (Wang W. Nat Rev Genet. 2007;8:735-748). Except in the case of FA-B patients (*FANCB* is located in the X chromosome), FA is autosomal recessive.

[000109] Proteins encoded by FA genes participate in a biochemical route known as the FA/BRCA pathway (See Wang W. *Nat Rev Genet.* 2007;8:735-748). Thirteen FA proteins have been identified in the FA pathway, each of them participating in one of the three FA protein complexes characterized so far in this pathway. The upstream complex, the FA core complex, is integrated by eight FA proteins (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, FANCM) and two FA associated proteins (FAAP24 and FAAP100). A second complex is formed by FANCD2 and FANCI, which work together in the FA-ID complex. Due to the E3 ligase activity (FANCL) of the FA core complex, FANCD2 and FANCI can be mono-ubiquitinated and then loaded onto chromatin, forming large nuclear foci in response to DNA damage or replication arrest. Finally, mono-ubiquitinated FANCD2/FANCI interact with downstream FA proteins such as FANCF/BRIP1, FANCG/PALB2 and FANCD1/BRCA2, which form stable complexes with proteins participating in homology directed repair (HDR), like BRCA1 and RAD51.

[000110] FA-A is the most frequent FA complementation group with about 50%-80% of FA patients corresponding to this complementation group (Casado JA *et al.* *J Med Genet.* 2007;44:241-249; Levitus M *et al.* *Blood.* 2004;103:2498-2503; Taniguchi T, D'andrea AD. *Blood.* 2006;107:4223-4233). As a consequence of the deficient or null expression of FANCA, the FA core complex cannot be formed. This prevents the activation of the ID complex, and consequently the migration of these proteins to chromatin, thus resulting in the characteristic phenotype of FA cells.

[000111] As reviewed by D'Andrea *et al.* (*Blood.* 1997;90:1725-1736), FA cells are characterized by different cellular phenotypes, mainly related to defects in cell survival, DNA repair and genomic stability.

[000112] FA is mainly characterized by congenital abnormalities, development of bone marrow failure, and a high risk of developing acute myeloid leukemia and certain solid tumors. On average, 70% of FA patients have congenital defects. The skeletal abnormalities (radial ray, hip, vertebral scoliosis, rib), and generalized skin hyperpigmentation, café au lait spots, are present in 60-70% of FA patients. Most patients have short stature, and around one-third of them have microphthalmia and renal abnormalities. In about 30% of FA patients, no obvious congenital abnormalities are observed (Tischkowitz M, Dokal I. *Br J Haematol.* 2004;126:176-191).

[000113] The most important clinical features of FA patients are hematological. Bone marrow failure (BMF) is the main characteristic of the disease. It generally appears between the ages of 5 and 10 years. Eighty percent of 15 year-old patients develop BMF, with the actuarial risk of BMF above 90% by 40 years of age (Butturini *et al.*, 1994, Kutler *et al.*, 2003). Thrombocytopenia or leukopenia typically precedes anemia. Pancytopenia generally worsens over time. Neutropenia is associated with an increased risk for infections.

[000114] FA patients are also prone to develop cancer, principally, acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS). The majority of tumors associated with FA develop after age 13 years, with an average age of 23 years. The relative risk for AML is increased 785-fold, with the median age of FA patients who develop AML being 14 years, and the cumulative incidence of hematological malignancy 30-55% by 40 years of age (Kutler *et al.*, 2003; Rosenberg PS *et al.* Blood. 2003;101:822-826). Older FA patients also have a high risk to develop solid tumors, mainly squamous cell carcinomas (SCC). The median age at which these patients develop solid tumors is 26 years, being the cumulative incidence of solid tumors 30% by the age of 40 (Kutler *et al.*, 2003, Rosenberg *et al.*, 2003).

[000115] Due to the complex clinical manifestations of FA, management of these patients is mainly focused on reducing symptoms of bone marrow failure (BMF), myeloid leukemia, and solid tumors. The limitations of each of the current therapies for patients with FA are reported in orphan medicinal products documentation for the lentiviral vector carrying the FANCA gene, whose sponsor is CIEMAT/CIBER on Rare Diseases (Bueren, J. (2010). Center for Biomedical Network Research on Rare Diseases Ref EU/3/10/822).

[000116] In certain embodiments, methods and compositions are provided for preparation of gene therapy vector compositions, e.g., viral vectors, comprising these genetic expression cassettes for use in the preparation of medicaments useful in central and targeted gene therapy of diseases, disorders, and dysfunctions in an animal, and in humans in particular.

[000117] In some embodiments, the present invention provides for gene therapy for Fanconi Anemia based on a LV vector harbouring the hPGK eukaryotic promoter that drives the expression of the *FANCA* cDNA. This therapeutic vector may be used to transduce human hematopoietic stem cells (HSCs), which may be subsequently transplanted into humans with Fanconi Anemia.

[000118] In certain embodiments, the present invention provides an FANCA LV vector for the genetic correction of Fanconi Anemia. Overall, results demonstrate feasibility of gene therapy for FA with a LV designed for clinical application.

[000119] The present disclosure includes gene expression cassettes (e.g., therapeutic cassettes), gene transfer cassettes comprising the gene expression cassettes (e.g., integration cassettes), plasmids comprising the gene transfer cassettes, and gene delivery vectors comprising the gene transfer cassettes. The gene expression cassettes, gene transfer cassettes, plasmids and gene delivery vectors comprising a polynucleotide sequence encoding a therapeutic gene product operably linked to a promoter sequence. In certain embodiments, the polynucleotide sequence is DNA or RNA. In certain embodiments, the gene expression cassette is a polynucleotide, the gene transfer cassette is a polynucleotide, and the vector is a virus, e.g., a lentivirus.

[000120] In certain embodiments, the therapeutic gene product is a FANCA protein or a functional fragment or variant thereof, optionally a wild-type human FANCA protein.

[000121] In particular embodiments, a gene expression cassette comprises a promoter region, a coding sequence, and a post-transcriptional regulatory element. In certain embodiments, the promoter region comprises a promoter sequence, or a functional fragment thereof. In one embodiment, the promoter is a human PGK promoter. In some embodiments, the expression cassette also comprises an RNA export signal. The RNA export signal may comprise a wPRE sequence. In some embodiments, a mutated wPRE, lacking any residual open reading frame (Schambach, Bohne et al. 2006) is included to improve the level of expression and stability of the therapeutic gene.

[000122] Some embodiments of the present invention comprise gene expression cassettes for the enhanced expression of a FANCA gene product. In some embodiments, the polynucleotide cassette comprises a wild type *FANCA* cDNA coding sequence, or a codon-optimized version of the human *FANCA* cDNA to increase mRNA stability upon transcription. For the optimization, GeneArt[®] software may be used, increasing the GC content and removing cryptic splice sites in order to avoid transcriptional silencing and therefore increase transgene expression. Alternatively, any optimization method known in the art may be used.

[000123] In some aspects of the invention, pharmaceutical compositions are provided comprising a gene delivery vector of the invention and a pharmaceutical excipient. In some embodiments, the pharmaceutical composition comprises a gene delivery vector of the invention and a pharmaceutical excipient.

[000124] In some aspects of the invention, methods are provided for expressing a transgene in mammalian cells. In some embodiments, the method comprises contacting one or more mammalian cells with an effective amount of a polynucleotide cassette of the invention or a gene delivery vector of the invention, wherein the transgene is expressed at detectable levels in the one or more mammalian cells. In some embodiments, the method comprises contacting one or more mammalian cells with an effective amount of a polynucleotide cassette of the invention or a gene delivery vector of the invention, wherein the transgene is expressed at therapeutic levels in the one or more mammalian cells. In some embodiments, the method is in vitro. In other embodiments, the method is in vivo.

[000125] In some aspects of the invention, methods are provided for the treatment or prophylaxis of a disease or disorder in a mammal in need of treatment or prophylaxis for a disease or disorder. In some embodiments, the method comprises administering to the mammal an effective amount of a pharmaceutical composition of the invention, wherein the coding sequence encodes a therapeutic gene product.

COMPOSITIONS

[000126] In some aspects of the disclosure, compositions are provided for the expression of a FANCA transgene in eukaryotic cells. In certain embodiments, the eukaryotic cell is a mammalian cell. In one embodiment, the mammalian cell is a hematopoietic stem cell (HSC). In one embodiment, the mammalian cell is a hematopoietic progenitor. In one embodiment, the mammalian cell is CD34⁺. In one embodiment, the mammalian cell is a

human cell. In particular embodiments, the cell is a human CD34⁺ cell derived from a subject diagnosed with FA who is to be treated with a the CD34⁺ cell after it is transduced with a gene delivery disclosed herein, and comprises a gene expression cassette disclosed.

[000127] In one specific embodiment, the present disclosure includes a lentiviral vector comprising a gene expression cassette comprising a polynucleotide sequence encoding a therapeutic FANCA protein or a functional fragment or variant thereof. As used herein, a functional variant of a reference polynucleotide or polypeptide comprises one or more amino acid or nucleic acid deletions, additions or substitutions, as compared to the reference sequence, and it retains at least 50%, at least 80%, at least 90%, or at least 99% of the functional activity of the reference polynucleotide or polypeptide. As used herein, a functional fragment is a fragment of a reference polynucleotide or polypeptide, and it retains at least 50%, at least 80%, at least 90%, or at least 99% of the functional activity of the reference polynucleotide or polypeptide.

[000128] In one embodiment, the backbone of the lentiviral vector is the same as the one corresponding to the medicinal product “lentiviral vector carrying the Wiscott Aldrich Syndrome Protein (WASP-LV)” (Ref141/2000), although for FA treatment, the promoter is the human phosphoglycerate kinase (hPGK) promoter, characterized by its stable activity *in vivo* and by improved safety properties, compared to other promoters already used in gene therapy (Modlich U, Navarro S, Zychlinski D *et al.* Insertional Transformation of Hematopoietic Cells by Self-Inactivating Lentiviral And Gammaretroviral Vectors. Mol Ther. 2009;17:1919- 1928; Montini E, Cesana D, Schmidt M *et al.* Hematopoietic Stem Cell Gene Transfer in a Tumor prone Mouse Model Uncovers Low Genotoxicity Of Lentiviral Vector Integration. Nat Biotechnol. 2006;24:687-696.).

[000129] In some embodiments of the disclosure, the composition comprises a polynucleotide cassette. By a “polynucleotide cassette” is meant a polynucleotide sequence comprising two or more functional polynucleotide sequences, *e.g.*, regulatory elements, translation initiation sequences, coding sequences, termination sequences, etc., typically in operable linkage to one another. Likewise, by a “polynucleotide cassette for the expression of a transgene in a mammalian cell,” it is meant a combination of two or more functional polynucleotide sequences, *e.g.*, promoter, enhancer, 5'UTR, translation initiation sequence, coding sequence, termination sequences, etc. that promotes the expression of the transgene

in a cell. Gene expression cassettes and gene transfer cassettes are examples of polynucleotide cassettes.

[000130] In some embodiments, the polynucleotide cassettes of the present disclosure provide for enhanced expression of a transgene in mammalian cells. As demonstrated by the working examples of the present disclosure, the present inventors have discovered a number of polynucleotide elements, *i.e.*, improved elements as compared to those known in the art, which individually and synergistically provide for the enhanced expression of transgenes in mammalian cells. In certain embodiments, the arrangement of the two or more functional polynucleotide sequences within the polynucleotide cassettes of the present disclosure provide for enhanced expression of a transgene in mammalian cells. By “enhanced” it is meant that expression of the transgene is increased, augmented, or stronger, in cells carrying the polynucleotide cassettes of the present disclosure relative to in cells carrying the transgene operably linked to comparable regulatory elements, *e.g.*, as known in the art. Put another way, expression of the transgene is increased, augmented, or stronger, from the polynucleotide cassettes of the present disclosure relative to expression from a polynucleotide cassette not comprising the one or more optimized elements of the present disclosure, *i.e.*, a reference control. In certain embodiment, the enhanced expression is specific for or limited to one or more desired cell types.

[000131] For example, expression of the transgene may be enhanced, or augmented, or stronger, in cells comprising a polynucleotide cassette comprising a promoter disclosed herein than in cells that carry the transgene operably linked to a different promoter, *e.g.*, as known in the art. As another example, expression of the transgene may be enhanced, or increased, augmented, or stronger, in cells comprising a polynucleotide cassette comprising an enhancer sequence disclosed herein than in cells that carry the transgene operably linked to a different enhancer sequence.

[000132] Without wishing to be bound by theory, enhanced expression of a transgene in cells is believed to be due to a faster build-up of gene product in the cells or a more stable gene product in the cells. Thus, enhanced expression of a transgene by the polynucleotide cassettes of the subject disclosure may be observed in a number of ways. For example, enhanced expression may be observed by detecting the expression of the transgene following contact of the polynucleotide cassette to the cells sooner, *e.g.* 2 days sooner, 7 days sooner, 2 weeks sooner, 3 weeks sooner, 4 weeks sooner, 8 weeks sooner, 12 weeks

sooner or more, than expression would be detected if the transgene were operably linked to comparable regulatory elements, e.g., as known in the art. Enhanced expression may also be observed as an increase in the amount of gene product per cell. For example, there may be a 2-fold increase or more, e.g. a 3-fold increase or more, a 4-fold increase or more, a 5-fold increase or more, or a 10-fold increase or more in the amount of gene product per mammalian cell. Enhanced expression may also be observed as an increase in the number of mammalian cells that express detectable levels of the transgene carried by the polynucleotide cassette. For example, there may be a 2-fold increase or more, e.g. a 3-fold increase or more, a 4-fold increase or more, a 5-fold increase or more, or a 10-fold increase or more in the number of mammalian cells that express detectable levels of the transgene.

[000133] As another example, the polynucleotide of the present invention may promote detectable levels of the transgene in a greater percentage of cells as compared to a conventional polynucleotide cassette; for example, where a conventional cassette may promote detectable levels of transgene expression in, for example, less than 5% of the cells in a certain region, the polynucleotide of the present invention promotes detectable levels of expression in 5% or more of the cells in that region; e.g. 10% or more, 15% or more, 20% or more, 25% or more, 30% or more, 35% or more, 40% or more, or 45% or more, in some instances 50% or more, 55% or more, 60% or more, 65% or more, 70% or more, or 75% or more, for example 80% or more, 85% or more, 90% or more, or 95% or more of the cells that are contacted, will express detectable levels of gene product. Enhanced expression may also be observed as an alteration in the viability and/or function of the cells.

[000134] The polynucleotide cassettes of the present disclosure typically comprise a promoter region. Any suitable promoter region or promoter sequence therein can be used in the subject polynucleotide cassettes, so long as the promoter region promotes expression of a coding sequence in eukaryotic cells. In certain embodiments, the promoter region promotes expression of a coding sequence in mammalian cells. In some instances, the promoter is a ubiquitous promoter, i.e., it is a promoter that is active in a wide range of cells, tissues and species. In other instances, the promoter is a human PGK promoter.

[000135] Promoter and enhancer elements can be tissue specific or stage-specific. For example, a tissue-specific promoter or enhancer preferentially drives expression (or a higher level of expression) in one or more particular cell type. Examples of cell types include but are not limited to: hematopoietic stem cells, long term hematopoietic stem cells, short term

hematopoietic stem cells, multipotent progenitors, hematopoietic CD34+ cells and any cluster differentiation subpopulation within the CD34+ population. A stage-specific promoter or enhancer preferentially drives expression (or higher level of expression) during one or more specific stages of the cell cycle or development. These include but are not limited to beta-globin locus control region, spectrin promoter, and an erythroid specific promoter.

[000136] In some embodiments, the polynucleotide comprises one or more enhancers. Enhancers are nucleic acid elements known in the art to enhance transcription, and can be located anywhere in association with the gene they regulate, e.g. upstream, downstream, within an intron, etc. Any enhancer element can be used in the polynucleotide cassettes and gene therapy vectors of the present disclosure, so long as it enhances expression of the gene when used in combination with the promoter.

[000137] The coding sequence to be expressed in the cells can be any polynucleotide sequence, e.g. gene or cDNA that encodes a gene product, e.g. a polypeptide or RNA-based therapeutic (siRNA, antisense, ribozyme, shRNA, etc.). The coding sequence may be heterologous to the promoter sequence to which it is operably linked, i.e. not naturally operably associated with it. Alternatively, the coding sequence may be endogenous to the promoter sequence to which it is operably linked, i.e. is associated in nature with that promoter. The gene product may act intrinsically in the mammalian cell, or it may act extrinsically, e.g., it may be secreted. For example, when the transgene is a therapeutic gene, the coding sequence may be any gene that encodes a desired gene product or functional fragment or variant thereof that can be used as a therapeutic for treating a disease or disorder. In various preferred embodiments, the transgene encodes human FANCA, i.e. SEQ ID NO: 25.

[000138] In one embodiment of the invention, the transgene coding sequence is modified, or “codon optimized” to enhance expression by replacing infrequently represented codons with more frequently represented codons. The coding sequence is the portion of the mRNA sequence that encodes the amino acids for translation. During translation, each of 61 trinucleotide codons are translated to one of 20 amino acids, leading to a degeneracy, or redundancy, in the genetic code. However, different cell types, and different animal species, utilize tRNAs (each bearing an anticodon) coding for the same amino acids at different frequencies. When a gene sequence contains codons that are infrequently represented by the

corresponding tRNA, the ribosome translation machinery may slow, impeding efficient translation. Expression can be improved via “codon optimization” for a particular species, where the coding sequence is altered to encode the same protein sequence, but utilizing codons that are highly represented, and/or utilized by highly expressed human proteins (Cid-Arregui et al., 2003; J. Virol. 77: 4928). In one aspect of the present invention, the coding sequence of the transgene is modified to replace codons infrequently expressed in mammal or in primates with codons frequently expressed in primates. For example, in some embodiments, the coding sequence encoded by the transgene encodes a polypeptide having at least 85% sequence identity to a polypeptide encoded by a sequence disclosed above or herein, for example at least 90% sequence identity, e.g. at least 95% sequence identity, at least 98% identity, at least 99% identity, wherein at least one codon of the coding sequence has a higher tRNA frequency in humans than the corresponding codon in the sequence disclosed above or herein.

[000139] In an additional embodiment of the invention, the transgene coding sequence is modified to enhance expression by termination or removal of open reading frames (ORFs) that do not encode the desired transgene. An open reading frame (ORF) is the nucleic acid sequence that follows a start codon and does not contain a stop codon. ORFs may be in the forward or reverse orientation, and may be “in frame” or “out of frame” compared with the gene of interest. Such open reading frames have the potential to be expressed in an expression cassette alongside the gene of interest, and could lead to undesired adverse effects. In one aspect of the present invention, the coding sequence of the transgene has been modified to remove open reading frames by further altering codon usage. This was done by eliminating start codons (ATG) and introducing stop codons (TAG, TAA, or TGA) in reverse orientation or out-of-frame ORFs, while preserving the amino acid sequence and maintaining highly utilized codons in the gene of interest (i.e., avoiding codons with frequency < 20%). In the present invention, the transgene coding sequence may be optimized by either of codon optimization and removal of non-transgene ORFs or using both techniques. As will be apparent to one of ordinary skill in the art, it is preferable to remove or minimize non-transgene ORFs after codon optimization in order to remove ORFs introduced during codon optimization.

[000140] In some embodiments, a polynucleotide cassette comprises:

[000141] (i) a phosphoglycerate kinase (PGK) promoter sequence or a functional variant or fragment thereof;

[000142] (ii) a sequence encoding a human FANCA protein or a functional fragment or variant thereof; and:

[000143] (iii) a post-transcriptional regulatory element of the woodchuck hepatitis virus (WPRE) sequence.

[000144] In some embodiments, a polynucleotide cassette comprises:

[000145] (i) a human phosphoglycerate kinase (PGK) promoter sequence;

[000146] (ii) a sequence encoding a human FANCA protein; and:

[000147] (iii) a mutant WPRE sequence.

[000148] In some embodiments, a polynucleotide cassette comprises:

[000149] a) a 5' LTR, optionally a modified 5' LTR;

[000150] b) a cPPT sequence;

[000151] c) PGK promoter sequence, optionally a human PGK promoter sequence;

[000152] d) a sequence encoding a human FANCA protein, optionally a cDNA sequence or a codon optimized sequence;

[000153] e) a mutant wPRE sequence; and

[000154] f) a 3' LTR, optionally a modified 3' LTR.

[000155] In one embodiment, the modified WPRE is referred to as WPRE*. WPRE* is a modified WPRE that lacks an open reading frame (see, *e.g.*, Schambach et al, 2006 Gene Ther. 13:641-645).

[000156] In certain embodiments, a gene transfer cassette comprises one or more additional elements, *e.g.*, one or more elements selected from the following: 5' LTR, 3'LTR, cPPT, CTS, RRE, enhancer sequences, and packaging signals.

[000157] The RRE sequence improves the efficiency of gene transfer. In particular embodiments of any of the expression cassettes and gene delivery vectors described herein, the RRE sequence comprises or consists of any of the following sequences, or sequences having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequences:

(AGGAGCTTTGTTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGT
CAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAG
AACAATTTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGG

GGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCA
ACAGCTCCT (SEQ ID NO: 1); or

a sequence comprising or consisting of nucleotides 2649-2882 or SEQ ID NO:24.

[000158] The retroviral leader region contains the packaging signal (Ψ), which is involved in packaging the retroviral genome into the viral capsid. LV vectors were thought to require approximately 300 bp of the *Gag* gene in this region. Currently, this *Gag* sequence has been reduced to just 40 bp (Figure 65). In particular embodiments of any of the expression cassettes and gene delivery vectors described herein, the ψ sequence is an HIV-1 ψ sequence or the ψ sequence comprises or consists of any of the following sequences, or sequences having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequences:

CTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCG
GCGACTGGTGAGTACGCCAAAAATTTGACTAGCGGAGGCTAGAAGGAGAGAGATG
GGTGCGAGAGCGTC (SEQ ID NO: 2); or

a sequence comprising or consisting of polynucleotides 2031-2156 of SEQ ID NO:24.

[000159] In particular embodiments of any of the expression cassettes and gene delivery vectors described herein, the truncated HIV-1 5' LTR comprises or consists of any of the following sequences, or sequences having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to any of the following sequences:

GGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAAC
CCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCCGT
CTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAA
ATCTCTAGCA (SEQ ID NO: 3); or

a sequence comprising or consisting of polynucleotides 1586-9495 of SEQ ID NO:24.

[000160] In particular embodiments of any of the expression cassettes and gene delivery vectors described herein, the HIV-1 self-inactivating 3' LTR comprises or consists of any of the following sequences, or sequences having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequences:

TGGAAGGGCTAATCACTCCCAACGAAGACAAGATCTGCTTTTTGCTTGTACTGGGT
CTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCAC
TGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCCGTCTGTT

GTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTC
TAGCA (SEQ ID NO: 4); or

a sequence comprising or consisting of polynucleotides 9262-9495 of SEQ ID NO:24.

[000161] In particular embodiments of any of the expression cassettes and gene delivery vectors described herein, the human cytomegalovirus (CMV) immediate early promoter comprises or consists of any of the following sequences, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to any of the following sequences:

GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA
TTTCCAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA
CGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCATGACGCAAATGGGCGGTAG
GCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT (SEQ ID NO: 5); or

a sequence comprising or consisting of polynucleotides 1586-1789 of SEQ ID NO:24.

[000162] The cPPT, which facilitates nuclear translocation of the pre-integration complexes, together with the CTS involved in the separation of reverse transcriptase, has been seen to improve viral titer (Zennou, et al. 2000; Follenzi et al. 2000). In particular embodiments of any of the expression cassettes and gene delivery vectors described herein, the central polypurine tract and central termination sequence of HIV-1 (cPPT/CTS) comprises or consists of any of the following sequences, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to any of the following sequences:

TTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGAC
ATAATAGCAACAGACATACAACTAAAGAATTACAAAAACAAATTACAAAAATTCA
AAATTTT (SEQ ID NO: 6);

TTTAAAAGAAAAGGGGGGATTGGGGGGT (SEQ ID NO:12); or

a sequence comprising or consisting of nucleotides 3378-3495 of SEQ ID NO:24.

[000163] In particular embodiments of any of the expression cassettes and gene delivery vectors described herein, the human phosphoglycerate kinase 1 (hPGK) promoter comprises or consists of any of the following sequences, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to any of the following sequences:

GGGGTTGGGGTTGCGCCTTTTCCAAGGCAGCCCTGGGTTTGCGCAGGGACGCGGCTG

CTCTGGGCGTGGTTCCGGGAAACGCAGCGGCGCCGACCCTGGGTCTCGCACATTCTT
CACGTCCGTTTCGCAGCGTCACCCGGATCTTCGCCGCTACCCTTGTGGGCCCCCGGC
GACGCTTCCTGCTCCGCCCCTAAGTCGGGAAGGTTTCCTTGCGGTTTCGCGGCGTGCCG
GACGTGACAAACGGAAGCCGCACGTCTCACTAGTACCCTCGCAGACGGACAGCGCC
AGGGAGCAATGGCAGCGCGCCGACCGCGATGGGCTGTGGCCAATAGCGGCTGCTCA
GCAGGGCGCGCCGAGAGCAGCGGCCGGAAGGGGCGGTGCGGGAGGCGGGGTGTG
GGGCGGTAGTGTGGGCCCTGTTCTGCCCCGCGCGGTGTTCCGCATTCTGCAAGCCTC
CGGAGCGCACGTGCGCAGTCGGCTCCCTCGTTGACCGAATCACCGACCTCTCTCCCC
AG (SEQ ID NO: 7); or
a sequence comprising or consisting of nucleotides 3541-4051 of SEQ ID NO:24.

[000164] Since most FA patients belong to the FA-A complementation group (Casado *et al.*, 2007, Levitus *et al.*, 2004, Taniguchi *et al.*, 2006), in particular embodiments, the encoded therapeutic gene product is FANCA, although the disclosure contemplates that FA proteins of other complementation groups may also be delivered, and thus encoded in the expression cassettes disclosed herein, e.g., instead of FANCA.

[000165] In particular embodiments of any of the expression cassettes and gene delivery vectors described herein, the polynucleotide sequence encoding FANCA is a human *FANCA* cDNA sequence that comprises or consists of the following sequence, or a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence:

ATGTCCGACTCGTGGGTCCCGAACTCCGCCTCGGGCCAGGACCCAGGGGGCCGC
CGGAGGGCCTGGGCCGAGCTGCTGGCGGGAAGGGTCAAGAGGGAAAAATATAA
TCCTGAAAGGGCACAGAAATTAAAGGAATCAGCTGTGCGCCTCCTGCGAAGCC
ATCAGGACCTGAATGCCCTTTTGCTTGAGGTAGAAGGTCCACTGTGTAAAAAAT
TGTCTCTCAGCAAAGTGATTGACTGTGACAGTTCTGAGGCCTATGCTAATCATTC
TAGTTCATTTATAGGCTCTGCTTTGCAGGATCAAGCCTCAAGGCTGGGGGTTCCT
GTGGGTATTCTCTCAGCCGGGATGGTTGCCTCTAGCGTGGGACAGATCTGCACG
GCTCCAGCGGAGACCAGTCACCCCTGTGCTGCTGACTGTGGAGCAGAGAAAGAA
GCTGTCTTCCCTGTTAGAGTTTGCTCAGTATTTATTGGCACACAGTATGTTCTCC
CGTCTTTCCTTCTGTCAAGAATTATGGAAAATACAGAGTTCTTTGTTGCTTGAAG
CGGTGTGGCATCTTCACGTACAAGGCATTGTGAGCCTGCAAGAGCTGCTGGAAA

GCCATCCCGACATGCATGCTGTGGGATCGTGGCTCTTCAGGAATCTGTGCTGCCT
TTGTGAACAGATGGAAGCATCCTGCCAGCATGCTGACGTCGCCAGGGCCATGCT
TTCTGATTTTGTTCAAATGTTTGTGTTTGTAGAGGGATTTCAGAAAACTCAGATCTG
AGAAGAACTGTGGAGCCTGAAAAAATGCCGCAGGTCACGGTTGATGTACTGCA
GAGAATGCTGATTTTTGCACTTGACGCTTTGGCTGCTGGAGTACAGGAGGAGTC
CTCCACTCACAAGATCGTGAGGTGCTGGTTCGGAGTGTTCAAGTGGACACACGCT
TGGCAGTGTAATTTCCACAGATCCTCTGAAGAGGTTCTTCAGTCATACCCTGACT
CAGATACTCACTCACAGCCCTGTGCTGAAAGCATCTGATGCTGTTCAAGATGCAG
AGAGAGTGGAGCTTTGCGCGGACACACCCTCTGCTCACCTCACTGTACCGCAGG
CTCTTTGTGATGCTGAGTGCAGAGGAGTTGGTTGGCCATTTGCAAGAAGTTCTG
GAAACGCAGGAGGTTCACTGGCAGAGAGTGCTCTCCTTTGTGTCTGCCCTGGTT
GTCTGCTTTCCAGAAGCGCAGCAGCTGCTTGAAGACTGGGTGGCGCGTTTGATG
GCCCAGGCATTCGAGAGCTGCCAGCTGGACAGCATGGTCACTGCGTTCCTGGTT
GTGCGCCAGGCAGCACTGGAGGGCCCCTCTGCGTTCCTGTTCATATGCAGACTGG
TTCAAGGCCTCCTTTGGGAGCACACGAGGCTACCATGGCTGCAGCAAGAAGGCC
CTGGTCTTCCTGTTTACGTTCTTGTCAGAACTCGTGCCTTTTGAGTCTCCCCGGTA
CCTGCAGGTGCACATTCTCCACCCACCCCTGGTTCCCAGCAAGTACCGCTCCCTC
CTCACAGACTACATCTCATTGGCCAAGACACGGCTGGCCGACCTCAAGGTTTCT
ATAGAAAACATGGGACTCTACGAGGATTTGTCATCAGCTGGGGACATTACTGAG
CCCCACAGCCAAGCTCTTCAGGATGTTGAAAAGGCCATCATGGTGTTTGAGCAT
ACGGGGAACATCCAGTCACCGTCATGGAGGCCAGCATATTCAGGAGGCCTTAC
TACGTGTCCCACTTCCTCCCCGCCCTGCTCACACCTCGAGTGCTCCCCAAAGTCC
CTGACTCCCGTGTGGCGTTTATAGAGTCTCTGAAGAGAGCAGATAAAATCCCCC
CATCTCTGTACTCCACCTACTGCCAGGCCTGCTCTGCTGCTGAAGAGAAGCCAG
AAGATGCAGCCCTGGGAGTGAGGGCAGAACCCAACTCTGCTGAGGAGCCCCCTG
GGACAGCTCACAGCTGCACTGGGAGAGCTGAGAGCCTCCATGACAGACCCCAG
CCAGCGTGATGTTATATCGGCACAGGTGGCAGTGATTTCTGAAAGACTGAGGGC
TGTCTGGGCCACAATGAGGATGACAGCAGCGTTGAGATATCAAAGATTCACT
CAGCATCAACACGCCGAGACTGGAGCCACGGGAACACATTGCTGTGGACCTCCT
GCTGACGTCTTTCTGTCAGAACCTGATGGCTGCCTCCAGTGTCGCTCCCCCGGAG
AGGCAGGGTCCCTGGGCTGCCCTCTTCGTGAGGACCATGTGTGGACGTGTGCTC
CCTGCAGTGCTACCCGGCTCTGCCAGCTGCTCCGTCACCAGGGCCCCGAGCCTG

AGTGCCCCACATGTGCTGGGGTTGGCTGCCCTGGCCGTGCACCTGGGTGAGTCC
AGGTCTGCGCTCCCAGAGGTGGATGTGGGTCCTCCTGCACCTGGTGCTGGCCTT
CCTGTCCCTGCGCTCTTTGACAGCCTCCTGACCTGTAGGACGAGGGATTCTTGT
TCTTCTGCCTGAAATTTTGTACAGCAGCAATTTCTTACTCTCTCTGCAAGTTTTCT
TCCCAGTCACGAGATACTTTGTGCAGCTGCTTATCTCCAGGCCTTATTA AAAAGT
TTCAGTTCCTCATGTTT CAGATTGTTCTCAGAGGCCCGACAGCCTCTTTCTGAGGA
GGACGTAGCCAGCCTTTCCTGGAGACCCTTGACCTTCCTTCTGCAGACTGGCA
GAGAGCTGCCCTCTCTCTCTGGACACACAGAACCTTCCGAGAGGTGTTGAAAGA
GGAAGATGTTCACTTAACTTACCAAGACTGGTTACACCTGGAGCTGGAAATTCA
ACCTGAAGCTGATGCTCTTTCAGATACTGAACGGCAGGACTTCCACCAGTGGGC
GATCCATGAGCACTTTCTCCCTGAGTCCTCGGCTTCAGGGGGCTGTGACGGAGA
CCTGCAGGCTGCGTGTACCATTTCTTGTC AACGCACTGATGGATTTCCACCAAAG
CTCAAGGAGTTATGACCACTCAGAAAATTCTGATTTGGTCTTTGGTGGCCGCAC
AGGAAATGAGGATATTATTTCCAGATTGCAGGAGATGGTAGCTGACCTGGAGCT
GCAGCAAGACCTCATAGTGCCTCTCGGCCACACCCCTTCCCAGGAGCACTTCCT
CTTTGAGATTTTCCGCAGACGGCTCCAGGCTCTGACAAGCGGGTGGAGCGTGGC
TGCCAGCCTTCAGAGACAGAGGGAGCTGCTAATGTACAAACGGATCCTCCTCCG
CCTGCCTTCGTCTGTCTCTGCGGCAGCAGCTTCCAGGCAGAACAGCCCATCACT
GCCAGATGCGAGCAGTTCTTCCACTTGGTCAACTCTGAGATGAGAACTTCTGC
TCCCACGGAGGTGCCCTGACACAGGACATCACTGCCCACTTCTTCAGGGGCCTC
CTGAACGCCTGTCTGCGGAGCAGAGACCCCTCCCTGATGGTCGACTTCATACTG
GCCAAGTGCCAGACGAAATGCCCCCTTAATTTTGACCTCTGCTCTGGTGTGGTGG
CCGAGCCTGGAGCCTGTGCTGCTCTGCCGGTGGAGGAGACACTGCCAGAGCCCG
CTGCCCCGGGA ACTGCAGAAGCTACAAGAAGGCCGGCAGTTTGCCAGCGATTTCT
CTCTCCCCTGAGGCTGCCTCCCCAGCACCCAACCCGGACTGGCTCTCAGCTGCTG
CACTGCACTTTGCGATTCAACAAGTCAGGGAAGAAAACATCAGGAAGCAGCTA
AAGAAGCTGGACTGCGAGAGAGAGGAGCTATTGGTTTTCTTTTCTTCTCTCCT
TGATGGGCCTGCTGTCGTCACATCTGACCTCAAATAGCACACAGACCTGCCAA
AGGCTTTCCACGTTTGTGCAGCAATCCTCGAGTGTTTAGAGAAGAGGAAGATAT
CCTGGCTGGCACTCTTTCAGTTGACAGAGAGTGACCTCAGGCTGGGGCGGCTCC
TCCTCCGTGTGGCCCCGGATCAGCACACCAGGCTGCTGCCTTTCGCTTTTTACAG
TCTTCTCTCCTACTTCCATGAAGACGCGGCCATCAGGGAAGAGGCCTTCCTGCAT

GTTGCTGTGGACATGTACTTGAAGCTGGTCCAGCTCTTCGTGGCTGGGGATACA
 AGCACAGTTTTCACCTCCAGCTGGCAGGAGCCTGGAGCTCAAGGGTCAGGGCAA
 CCCC GTGGA ACTGATAACAAAAGCTCGTCTTTTTCTGCTGCAGTTAATACCTCGG
 TGCCCGAAAAAGAGCTTCTCACACGTGGCAGAGCTGCTGGCTGATCGTGGGGAC
 TGCGACCCAGAGGTGAGCGCCGCCCTCCAGAGCAGACAGCAGGCTGCCCCTGA
 CGCTGACCTGTCCCAGGAGCCTCATCTCTTCTGA

(SEQ ID NO: 8).

[000166] The present disclosure includes plasmids comprising an expression cassette or transfer cassette described herein. In particular embodiments, the plasmid is pCCL-PGK-FANCA-WPRE* (Figure 41; SEQ ID NO: 24).

[000167] In certain embodiments, the disclosure includes a cell, e.g., a packaging cell or packaging cells line, e.g., 293 cells, comprising a plasmid disclosed herein. In particular embodiments, the cell comprises the plasmids depicted in Figures 38-41.

[000168] In certain embodiments, a transfer cassette or plasmid disclosed herein further comprises one or more additional elements, e.g., a CMV promoter and/or enhancer, an SV40 polyA sequence, an origin of replication, e.g., an SV40 ori sequence, or any of the elements disclosed herein.

[000169] In particular embodiments of any of the transfer cassettes, plasmids or vectors described herein, the human CMV enhancer comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence:
 GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
 TAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGG
 CTGACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCAT
 AGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTA
 AACTGCCC ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTAT
 TGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT
 ATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG
 (SEQ ID NO: 9).

[000170] In particular embodiments of any of the transfer cassettes, plasmids or vectors described herein, the simian virus 40 (SV40) poly(A) signal comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at

least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence:

AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAAT
TTCACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCA
TCAATGTATCTTA (SEQ ID NO: 10).

[000171] In particular embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the SV40 origin of replication comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence:

ATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAA
TTTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAA
GTAGTGAGGAGGCTTTTTTGGAGGCC (SEQ ID NO: 11).

[000172] In some embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the dNEF signal present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence:

GAATTTCGAGCTCGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCT
TAGCCACTTTTTTAAAAGAAAAGGGGGGAC (SEQ ID NO: 13).

[000173] In particular embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the KanR sequence present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence:

ATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCGGCTTGGGTGGAGAGG
CTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTG
TTCCGGCTGTCAGCGCAGGGGCGTCCGGTTCTTTTTGTCAAGACCGACCTGTCCG
GTGCCCTGAATGAACTGCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCGACG
ACGGGCGTTCTTGCGCGGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGAC
TGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTC
CTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTG
ATCCGGCTACCTGCCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGAGCAC

GTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATC
 AGGGGCTCGCGCCAGCCGAAGTTCGCCAGGCTCAAGGCGTCTATGCCCCGACG
 GCGAGGATCTCGTCGTGACCCACGGCGATGCCTGCTTGCCGAATATCATGGTGG
 AAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGTCTGGGTGTGGCGGACC
 GCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCG
 AATGGGCTGACCGCTTCCTTGTGCTTTACGGTATCGCCGCGCCCGATTTCGCAGCG
 CATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA (SEQ ID NO: 14)

[000174] In some embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the *rrnG* terminator (transcription terminator from the *E.coli* ribosomal RNA *rrnG* operon (Albrechtsen et al., 1991) present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence:

GCATTGGCGCAGAAAAAATGCCTGATGCGACGCTGCGCGTCTTATACTCCCAC
 ATATGCCAGATTCAGCAACGGATACGGCTTCCCCAACTTGCCCACTTCCATACG
 TGTCTCCTTACCAGAAATTTATCCTTAA (SEQ ID NO: 15)

[000175] In some embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the *ori* (high-copy-number ColE1/pMB1/pBR322/pUC origin of replication) present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence:

TTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCG
 CTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGG
 TAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGT
 AGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCT
 AATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTT
 GGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGG
 GTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACC
 TACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGAC
 AGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCC
 AGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCTGGGTTTCGCCACCTCTGACTT
 GAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAA (SEQ ID

NO: 16).

[000176] In some embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the CAP binding site present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence: TAATGTGAGTTAGCTCACTCAT (SEQ ID NO: 17).

[000177] In some embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the *E.coli lac* promoter present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence: TTTACACTTTATGCTTCCGGCTCGTATGTTG (SEQ ID NO: 18).

[000178] In some embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the lac operator present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence: TTGTGAGCGGATAACAA (SEQ ID NO: 19)

[000179] In some embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the T3 promoter (promoter for bacteriophage T3 RNA polymerase) present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence: AATTAACCCTCACTAAAGG (SEQ ID NO: 20).

[000180] In some embodiments of any of the transfer cassettes, plasmids or vectors described herein described herein, the T7 promoter (promoter for bacteriophage T7 RNA polymerase) present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at

least 99% identity to the following sequence: CCTATAGTGAGTCGTATTA (SEQ ID NO: 21).

[000181] In some embodiments of any of the transfer cassettes, plasmids or vectors described herein, the f1 ori (f1 bacteriophage origin of replication) present in any of the expression cassettes or gene delivery vectors described herein comprises or consists of the following sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence:

ACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTC
CTTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCT
TTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAG
GGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGA
CGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACT
CAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCT
ATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATT (SEQ ID NO: 22).

[000182] As discussed herein, the polynucleotide cassettes of the present invention may comprise an RNA export signal. Exemplary RNA export sequences include but are not limited to wPRE. The wPRE significantly increases transgene expression in target cells, by increasing RNA stability in a transgene, promoter and vector-independent manner (Zuffrey et al, 1999). However, it can express a truncated 60-amino acid protein derived from the WHV X gene involved in liver cancer (Kingsman et al, 2005). Therefore, most pre-clinical protocols and clinical trials include a mutated version of the wPRE element (Zanta-Boussif et al, 2009). On the other hand, the use of two SV40-USE elements in SIN-LV vectors has been seen to be more efficient than the wPRE sequence in suppressing transcriptional read through (Schambach et al, 2007). More precisely, the wPRE disclosed herein is a chimeric wPRE that carries 589 nucleotides from the modified WPRE performed by Axel Schambach (nucleotides 1-589) (WO 2008136670 A2; [5]) and 88 from a former wPRE (nucleotide 590-677) (Zuffrey et al, 1999). Data disclosed herein shows this chimeric wPRE works better than the former wPRE. The chimeric wPRE sequence comprises the following

sequence, a functional fragment thereof, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to the following sequence:

CGAGCATCTTACCGCCATTTATTTCCCATATTTGTTCTGTTTTTCTTGATTTGGGTA
TACATTTAAATGTTAATAAAACAAAATGGTGGGGCAATCATTACATTTTTAGG
GATATGTAATTACTAGTTCAGGTGTATTGCCACAAGACAAACATGTTAAGAAAC
TTTCCCGTTATTTACGCTCTGTTCCCTGTTAATCAACCTCTGGATTACAAAATTTGT
GAAAGATTGACTGATATTCTTAACCTATGTTGCTCCTTTTACGCTGTGTGGATATG
CTGCTTTAATGCCTCTGTATCATGCTATTGCTTCCCGTACGGCTTTTCGTTTTCTCC
TCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCC
GTCAACGTGGCGTGGTGTGCTCTGTGTTTGCTGACGCAACCCCCACTGGCTGGG
GCATTGCCACCACCTGTCAACTCCTTTCTGGGACTTTTCGCTTTCCCCCTCCCGAT
CGCCACGGCAGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTAG
GTTGCTGGGCACTGATAATTCCGTGGTGTGTCGGGGAAGGGCCTGCTGCCGGC
TCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTT
TGGGCCGCCTCCCCGCCTG (SEQ ID NO:23).

[000183] In particular embodiments, the mutated WPRE sequence comprises or consists of WPRE*, which corresponds to nucleotides 8502-9178 of SEQ ID NO:24, or has at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to this region of SEQ ID NO:24.

[000184] Other combinations of elements both as disclosed herein or as known in the art will be readily appreciated by the ordinarily skilled artisan.

[000185] Additionally, as will be recognized by one of ordinary skill in the art, the polynucleotide cassettes may optionally contain other elements including, but not limited to restriction sites to facilitate cloning and regulatory elements for a particular gene expression vector.

[000186] In some aspects of the present invention, the subject polynucleotide cassettes are used to deliver a gene to cells, e.g. to determine the effect that the gene has on cell viability and/or function, to treat a cell disorder, etc. In various embodiments, delivery of a viral vector to cells by transduction may occur in vitro, ex vivo, or in vitro. Accordingly, in some aspects of the invention, the composition that provides for the expression of a transgene in mammalian cells is a gene delivery vector, wherein the gene delivery vector comprises a polynucleotide cassette, e.g., a gene transfer cassette, of the present disclosure.

[000187] Any convenient gene delivery vector that finds use delivering polynucleotide sequences to mammalian cells is encompassed by the gene delivery vectors of the present disclosure. For example, the vector may comprise single or double stranded nucleic acid, e.g. single stranded or double stranded DNA. For example, the gene delivery vector may be DNA, e.g., a naked DNA, e.g., a plasmid, a minicircle, etc. The vector may comprise single-stranded or double-stranded RNA, including modified forms of RNA. In another example, the gene delivery vector may be an RNA, e.g., an mRNA or modified mRNA.

[000188] As another example, the gene delivery vector may be a viral vector derived from a virus, e.g., an adenovirus, an adeno-associated virus, a lentivirus (LV), a herpes virus, an alphavirus or a retrovirus, e.g., Moloney murine leukemia virus (M-MuLV), Moloney murine sarcoma virus (MoMSV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), gibbon ape leukemia virus (GaLV), feline leukemia virus (FLV), spumavirus, Friend murine leukemia virus, Murine Stem Cell Virus (MSCV) or Rous Sarcoma Virus (RSV). While embodiments encompassing the use of LV are described in greater detail below, it is expected that the ordinarily skilled artisan will appreciate that similar knowledge and skill in the art can be brought to bear on non-LV gene therapy vectors as well.

[000189] In some embodiments, the gene delivery vector is a self-limiting LV. In a specific embodiment of any of the expression cassettes and gene delivery vectors described herein, the transfer cassette is a pCCL-SIN-cPPT/CTS –hPGK-hFANCA-WPRE (Figure 41) of the disclosure comprises or consists of the following sequence, or a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% identity to SEQ ID NO: 24. SEQ ID NO: 24 corresponds to the pCCL-PGK-FANCA-WPRE* plasmid of Figure 41.

[000190] In one embodiment, a FANCA gene is delivered via a lentiviral vector (LV). The FANCA LVs described herein utilize a self-inactivating lentiviral vector (LV). In one embodiment, the FANCA LV comprises a promoter of the human phosphoglycerate (PGK) gene. The safety properties of this vector have been markedly improved, compared to the gamma-retroviral vectors already used in the clinics, which harbored strong viral promoters.

[000191] In certain embodiments, the lentiviral vector is PGK-FANCA.WPRE*LV, which comprises the gene transfer cassette depicted in Figure 1, comprising sequences disclosed in SEQ ID NO: 24. The PGK-FANCA-WPRE*LV gene expression cassette portion comprises the

human PGK promoter, the coding sequence for FANCA cDNA, and the WPRE*; and corresponds to nucleotides 3541 to 9178 of SEQ ID NO: 24. The PGK-FANCA-WPRE*LV transfer cassette portion comprises from about the 5' LTR (U5) to about the 3' LTR (U5) of the sequence shown in Figure 41. With respect to SEQ ID NO: 24, nucleotides 1586-1789 of SEQ ID NO: 24 comprise human CMV immediate early promoter. Nucleotides 2031-2156 of SEQ ID NO: 24 comprise HIV1 psi packaging signal. Nucleotides 2649-2882 of SEQ ID NO: 24 comprise HIV1 RRE element. Nucleotides 3378-3495 of SEQ ID NO: 24 comprise HIV cPPT/CTS element. Nucleotides 3541-4051 of SEQ ID NO: 24 comprise the hPGK promoter. Nucleotides 4078-8445 of SEQ ID NO: 24 comprise human FANCA-A cDNA. Nucleotides 8502-9178 of SEQ ID NO: 24 comprise mutated WPRE element. Nucleotides 9262-9495 of SEQ ID NO: 24 comprise the HIV delta U 3' LTR.

[000192] In yet another embodiment, the lentiviral vector contains the following elements:

- (i) the backbone of the lentiviral vector derived from the initial pCCLsin-cppt-hPGK-eGFP-WPRE (Dull et al, 1998 ; J.Virol 72 (11), 9873-9880). The pCCL backbone utilizes a heterologous CMV-HIV 5' LTR to obtain high levels of viral RNA transcription in the producer cells. Such heterologous LTR renders the construct independent from the need to use the HIV Tat protein for the production of the rHIV particles and it is therefore a safety feature. The U3 region of the 3' LTR contains a 400 bp deletion as described in (Zufferey et al J Virol, 1998) which confers self inactivating properties to the vector;
- (ii) the cDNA of the human *FANCA* gene (4368 bp GenBank accession number: X_99226 or as disclosed herein) encoding the FANCA protein (1455 AA) under control of the human PGK promoter. The promoter has already been characterized by its stable activity *in vivo* and by improved safety properties, compared to other promoters already used in gene therapy; and
- (iii) a mutated version of the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) that is deleted in the 3' region of a sequence coding for the X protein and any residual ORF of the described by Schambach et al (Gene therapy, 2006; 13, 641-645) or WPRE*.

[000193] Gene therapy vectors encapsulating the polynucleotide cassettes of the present disclosure may be produced using standard methodology. For example, in the case of LV virions, an LV expression vector according to the invention may be introduced into a producer cell, followed by introduction of an LV helper construct, where the helper construct includes LV coding regions capable of being expressed in the producer cell and which complement LV helper functions absent in the LV vector. This is followed by

introduction of helper virus and/or additional vectors into the producer cell, wherein the helper virus and/or additional vectors provide accessory functions capable of supporting efficient LV virus production. The producer cells are then cultured to produce LV. These steps are carried out using standard methodology. In particular embodiments, the plasmids depicted in Figures 38-41 are used to produce the gene delivery vectors.

[000194] Any suitable method for producing viral particles for delivery of the subject polynucleotide cassettes can be used, including but not limited to those described in the examples that follow. Any concentration of viral particles suitable to effectively transduce mammalian cells can be prepared for contacting mammalian cells in vitro or in vivo. For example, the viral particles may be formulated at a concentration of 10^8 vector genomes per mL or more, for example, 5×10^8 vector genomes per mL; 10^9 vector genomes per mL; 5×10^9 vector genomes per mL; 10^{10} vector genomes per mL; 5×10^{10} vector genomes per mL; 10^{11} vector genomes per mL; 5×10^{11} vector genomes per mL; 10^{12} vector genomes per mL; 5×10^{12} vector genomes per mL; 10^{13} vector genomes per mL; 1.5×10^{13} vector genomes per mL; 3×10^{13} vector genomes per mL; 5×10^{13} vector genomes per mL; 7.5×10^{13} vector genomes per mL; 9×10^{13} vector genomes per mL; 1×10^{14} vector genomes per mL, 5×10^{14} vector genomes per mL or more, but typically not more than 1×10^{15} vector genomes per mL.

[000195] In preparing the subject LV compositions, any host cells for producing LV virions may be employed, including, for example, mammalian cells (e.g. 293 cells), insect cells (e.g. SF9 cells), microorganisms and yeast. Host cells can also be packaging cells in which the LV rep and cap genes are stably maintained in the host cell or producer cells in which the LV vector genome is stably maintained and packaged. Exemplary packaging and producer cells are derived from SF-9, 293, A549 or HeLa cells. LV vectors are purified and formulated using standard techniques known in the art.

[000196] In certain embodiments, the present invention includes a cell comprising a gene expression cassette, gene transfer cassette, or gene delivery vector disclosed herein. In related embodiments, the cell is transduced with a gene delivery vector comprising an expression cassette disclosed herein or has an expression cassette disclosed herein integrated into the cell's genome.

[000197] In certain embodiments, the cell is a cell used to produce a viral gene delivery vector, e.g., a packaging cell.

[000198] In other embodiments, the cell is a cell to be delivered to a subject in order to provide to the subject the gene product encoded by the expression cassette. Thus, in certain embodiments, the cell is autologous to the subject to be treated or was obtained from the subject to be treated. In other embodiments, the cell is allogeneic to the subject to be treated or was obtained from a donor other than the subject to be treated. In particular embodiments, the cell is a mammalian cell, e.g., a human cell. In certain embodiments, the cell is a blood cell, an erythrocyte, a hematopoietic progenitor cell, a bone marrow cell, e.g., a lineage depleted bone marrow cell, a hematopoietic stem cell (e.g., CD34+) or a committed hematopoietic erythroid progenitor cell. In particular embodiments, the cell is a CD34+ cell obtained from a subject to be treated with the cell after it is transduced by a gene delivery vector disclosed herein. In particular embodiment, the cell is a CD34+ FA cell obtained from a subject diagnosed with FA.

[000199] The present invention includes pharmaceutical compositions comprising a polynucleotide cassette, gene delivery vector, or cell described herein and a pharmaceutically-acceptable carrier, diluent or excipient. The subject polynucleotide cassette, gene delivery vector, or cell can be combined with pharmaceutically-acceptable carriers, diluents and reagents useful in preparing a formulation that is generally safe, non-toxic, and desirable, and includes excipients that are acceptable for primate use. Such excipients can be solid, liquid, semisolid, or, in the case of an aerosol composition, gaseous. Examples of such excipients, carriers or diluents include, but are not limited to, water, saline, Ringer's solutions, dextrose solution, and 5% human serum albumin. Supplementary active compounds can also be incorporated into the formulations. Solutions or suspensions used for the formulations can include a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial compounds such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating compounds such as ethylenediaminetetraacetic acid (EDTA); buffers such as acetates, citrates or phosphates; detergents such as Tween 20 to prevent aggregation; and compounds for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. In particular embodiments, the pharmaceutical compositions are sterile.

[000200] Pharmaceutical compositions suitable for use in the present invention further include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion.

[000201] Sterile solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[000202] In one embodiment, the compositions are prepared with carriers that will protect the gene cassette or expression vector against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially.

[000203] It is especially advantageous to formulate oral, ocular or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

[000204] The pharmaceutical compositions can be included in a container, pack, or dispenser, e.g. syringe, e.g. a prefilled syringe, together with instructions for administration.

[000205] The pharmaceutical compositions of the invention encompass any pharmaceutically acceptable salts, esters, or salts of such esters, or any other compound

which, upon administration to an animal comprising a human, is capable of providing (directly or indirectly) the biologically active metabolite or residue thereof.

[000206] The term “pharmaceutically acceptable salt” refers to physiologically and pharmaceutically acceptable salts of the compounds of the invention: i.e., salts that retain the desired biological activity of the parent compound and do not impart undesired toxicological effects thereto. A variety of pharmaceutically acceptable salts are known in the art and described, e.g., in in “Remington’s Pharmaceutical Sciences”, 17th edition, Alfonso R. Gennaro (Ed.), Mark Publishing Company, Easton, PA, USA, 1985 (and more recent editions thereof), in the “Encyclopaedia of Pharmaceutical Technology”, 3rd edition, James Swarbrick (Ed.), Informa Healthcare USA (Inc.), NY, USA, 2007, and in J. Pharm. Sci. 66: 2 (1977). Also, for a review on suitable salts, see Handbook of Pharmaceutical Salts: Properties, Selection, and Use by Stahl and Wermuth (Wiley-VCH, 2002).

[000207] Pharmaceutically acceptable base addition salts are formed with metals or amines, such as alkali and alkaline earth metals or organic amines. Metals used as cations comprise sodium, potassium, magnesium, calcium, and the like. Amines comprise N-N'-dibenzylethylenediamine, chlorprocaine, choline, diethanolamine, dicyclohexylamine, ethylenediamine, N-methylglucamine, and procaine (see, for example, Berge et al., “Pharmaceutical Salts,” J. Pharma Sci., 1977, 66, 119). The base addition salts of said acidic compounds are prepared by contacting the free acid form with a sufficient amount of the desired base to produce the salt in the conventional manner. The free acid form may be regenerated by contacting the salt form with an acid and isolating the free acid in the conventional manner. The free acid forms differ from their respective salt forms somewhat in certain physical properties such as solubility in polar solvents, but otherwise the salts are equivalent to their respective free acid for purposes of the present invention.

[000208] The subject polynucleotide cassette, gene delivery vector, e.g., recombinant virus (virions), or cell (e.g., transduced with a gene delivery vector disclosed herein) can be incorporated into pharmaceutical compositions for administration to mammalian patients, particularly primates and more particularly humans. The subject polynucleotide cassette, gene delivery vector, e.g. virions, or cell can be formulated in nontoxic, inert, pharmaceutically acceptable aqueous carriers, preferably at a pH ranging from 3 to 8, more preferably ranging from 6 to 8. Such sterile compositions will comprise the vector or virion

containing the nucleic acid encoding the therapeutic molecule dissolved in an aqueous buffer having an acceptable pH upon reconstitution.

[000209] In some embodiments, the pharmaceutical composition provided herein comprise a therapeutically effective amount of a cell, vector or virion disclosed herein in admixture with a pharmaceutically acceptable carrier and/or excipient, for example saline, phosphate buffered saline, phosphate and amino acids, polymers, polyols, sugar, buffers, preservatives and other proteins. Exemplary amino acids, polymers and sugars and the like are octylphenoxy polyethoxy ethanol compounds, polyethylene glycol monostearate compounds, polyoxyethylene sorbitan fatty acid esters, sucrose, fructose, dextrose, maltose, glucose, mannitol, dextran, sorbitol, inositol, galactitol, xylitol, lactose, trehalose, bovine or human serum albumin, citrate, acetate, Ringer's and Hank's solutions, cysteine, arginine, carnitine, alanine, glycine, lysine, valine, leucine, polyvinylpyrrolidone, polyethylene and glycol. Preferably, this formulation is stable for at least six months at 4° C.

[000210] In some embodiments, the pharmaceutical composition provided herein comprises a buffer, such as phosphate buffered saline (PBS) or sodium phosphate/sodium sulfate, tris buffer, glycine buffer, sterile water and other buffers known to the ordinarily skilled artisan such as those described by Good et al. (1966) *Biochemistry* 5:467. The pH of the buffer in which the pharmaceutical composition comprising the tumor suppressor gene contained in the adenoviral vector delivery system, may be in the range of 6.5 to 7.75, preferably 7 to 7.5, and most preferably 7.2 to 7.4.

[000211] In certain embodiments, viral vectors may be formulated into any suitable unit dosage, including, without limitation, 1×10^8 vector genomes or more, for example, 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , or 1×10^{13} vector genomes or more, in certain instances, 1×10^{14} vector genomes, but usually no more than 4×10^{15} vector genomes. In some cases, the unit dosage is at most about 5×10^{15} vector genomes, e.g. 1×10^{14} vector genomes or less, for example 1×10^{13} , 1×10^{12} , 1×10^{11} , 1×10^{10} , or 1×10^9 vector genomes or less, in certain instances 1×10^8 vector genomes or less, and typically no less than 1×10^8 vector genomes. In some cases, the unit dosage is 1×10^{10} to 1×10^{11} vector genomes. In some cases, the unit dosage is 1×10^{10} to 3×10^{12} vector genomes. In some cases, the unit dosage is 1×10^9 to 3×10^{13} vector genomes. In some cases, the unit dosage is 1×10^8 to 3×10^{14} vector genomes. In one embodiment, the range is from about 5×10^{10} to about 1×10^{11} vector genomes. In some embodiments, the range is from about 1×10^9 to about 1×10^{10} vector genomes.

[000212] In some cases, the unit dosage of a pharmaceutical composition may be measured using multiplicity of infection (MOI). By MOI it is meant the ratio, or multiple, of vector or viral genomes to the cells to which the nucleic acid may be delivered. In some cases, the MOI may be 1×10^6 . In some cases, the MOI may be 1×10^5 - 1×10^7 . In some cases, the MOI may be 1×10^4 - 1×10^8 . In some cases, recombinant viruses of the disclosure are at least about 1×10^1 , 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , 1×10^{13} , 1×10^{14} , 1×10^{15} , 1×10^{16} , 1×10^{17} , and 1×10^{18} MOI. In some cases, recombinant viruses of this disclosure are 1×10^8 to 3×10^{14} MOI. In some cases, recombinant viruses of the disclosure are at most about 1×10^1 , 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , 1×10^{13} , 1×10^{14} , 1×10^{15} , 1×10^{16} , 1×10^{17} , and 1×10^{18} MOI. In some, embodiments the range is from about 20 to about 400 MOI.

[000213] In some aspects, the amount of pharmaceutical composition comprises about 1×10^8 to about 1×10^{15} recombinant viruses, about 1×10^9 to about 1×10^{14} recombinant viruses, about 1×10^{10} to about 1×10^{13} recombinant viruses, or about 1×10^{11} to about 3×10^{12} recombinant viruses.

METHODS

[000214] As discussed in more detail below, the subject polynucleotide cassettes and gene delivery vectors, referred to collectively herein as the “subject compositions”, find use in expressing a transgene, *e.g.*, *FANCA*, in cells of an animal, *e.g.*, a mammal or human. For example, the subject compositions may be used in research, *e.g.*, to determine the effect that the gene has on cell viability and/or function. As another example, the subject compositions may be used in medicine, *e.g.*, to treat a disorder such as FA. Thus, in some aspects of the invention, methods are provided for the expression of a gene in cells, the method comprising contacting cells with a composition of the present disclosure. In some embodiments, contacting occurs *in vitro*. In some embodiments, contacting occurs *in vivo*, *i.e.*, the subject composition is administered to a subject.

[000215] For instances in which mammalian cells are to be contacted *in vitro* or *in vivo* with a subject polynucleotide cassette or gene delivery vector comprising a subject polynucleotide cassette, the cells may be from any mammalian species, *e.g.*, rodent (*e.g.*, mice, rats, gerbils, squirrels), rabbit, feline, canine, goat, ovine, pig, equine, bovine, primate, human. Cells may be from established cell lines or they may be primary cells, where “primary cells”, “primary cell lines”, and “primary cultures” are used interchangeably herein

to refer to cells and cells cultures that have been derived from a subject and allowed to grow *in vitro* for a limited number of passages, *i.e.*, splittings, of the culture. For example, primary cultures are cultures that may have been passaged 0 times, 1 time, 2 times, 4 times, 5 times, 10 times, or 15 times, but not enough times go through the crisis stage. Typically, the primary cell lines of the present invention are maintained for fewer than 10 passages *in vitro*.

[000216] Embodiments of the present invention comprise mammalian cells (e.g., CD34+ cells) transduced with a viral delivery vector, e.g., a LV vector containing the human FANCA gene. In addition, the present invention includes a method of transducing a mammalian cell, e.g. a human hematopoietic stem cell or other cell described herein, comprising contacting the cell with a gene delivery vector, e.g., a LV vector, disclosed herein or comprising an expression cassette described herein. In certain embodiments, the cell was previously obtained from a subject to be treated, or from another donor. In particular embodiments, the subject was diagnosed with Fanconi Anemia, and the cell is transduced with a LV comprising an expression cassette encoding a FANCA coding region or cDNA. It is understood that the disclosed methods, e.g., those used to deliver a FANCA gene product, e.g., using a FANCA cDNA sequence, to a subject may also be used to treat Fanconi Anemia. In particular embodiments, the transduced cells are a population of cells obtained from a subject with FA, who is to be treated with the cells once they have been transduced. The cells may be obtained from bone marrow or blood. In certain embodiments the subject with FA is treated with agents to mobilize stem cells, then blood is drawn from the subject, red blood cells are removed, and CD34+ cells are selected. Following selection, the cells are then transduced. In particular embodiments, the transduced cells are stored or frozen before use, whereas in certain embodiments, they are provided to the subject immediately or shortly after they are transduced, e.g., within one hour, two hours, or four hours.

[000217] In certain embodiments, when transducing a cell with a gene delivery vector disclosed herein, the cells are contacted with the gene delivery vector for about 30 minutes, about 1 hour, about 1.5 hours, about 2 hours, about 2.5 hours, about 3 hours, about 3.5 hours, about 4 hours, about 5 hours, about 6 hours, about 7 hours, about 8 hours, about 12 hours, about 16 hours, about 18 hours, about 20 hours, about 24 hours, about 36 hours, about

48 hours, about 60 hours. In some embodiments, the cells are transduced for less than 60 hours, less than 48 hours, less than 36 hours, or less than 24 hours.

[000218] The subject polynucleotide cassette or gene delivery vector comprising a subject polynucleotide cassette may be provided to the subject cells one or more times, e.g. one time, twice, three times, or more than three times, and the cells allowed to incubate with the agent(s) for some amount of time following each contacting event e.g. 16-24 hours, after which time the media is replaced with fresh media and the cells are cultured further.

Contacting the cells may occur in any culture media and under any culture conditions that promote the survival of the cells. The culture may contain growth factors to which the cells are responsive. Growth factors, as defined herein, are molecules capable of promoting survival, growth and/or differentiation of cells, either in culture or in the intact tissue, through specific effects on a transmembrane receptor. Growth factors include polypeptides and non-polypeptide factors.

[000219] Typically, an effective amount of subject gene delivery vector or transduced cells comprising a subject polynucleotide cassette is provided to produce the expression of the transgene in cells. As discussed elsewhere herein, the effective amount may be readily determined empirically, e.g. by detecting the presence or levels of transgene gene product, by detecting an effect on the viability or function of the cells, etc. Typically, an effect amount of subject polynucleotide cassette or gene delivery vector comprising a subject polynucleotide cassette will promote greater expression of the transgene in cells than the same amount of a polynucleotide cassette as known in the art. Typically, expression will be enhanced 2-fold or more relative to the expression from a reference, or control, polynucleotide cassette e.g. as known in the art, for example 3-fold, 4-fold, or 5-fold or more, in some instances 10-fold, 20-fold or 50-fold or more, e.g. 100-fold.

[000220] For instances in which cells are to be contacted in vivo with a subject polynucleotide cassette or gene delivery vector comprising a subject polynucleotide cassette, the subject may be any mammal, e.g. rodent (e.g. mice, rats, gerbils), rabbit, feline, canine, goat, ovine, pig, equine, bovine, or primate. In a further preferred embodiment, the primate is a human. In a further embodiment, the cells are CD34+ cells.

[000221] The methods and compositions of the present disclosure find use, e.g., in the treatment of Fanconi Anemia.

[000222] In some embodiments, the subject method results in a therapeutic benefit, *e.g.*, preventing the development of a disorder, halting the progression of a disorder, reversing the progression of a disorder, etc. For example, in one embodiment, the disorder is BMF. In one embodiment, the disorder is thrombocytopenia. In another embodiment, the disorder is leukopenia. In one embodiment, the disorder is pancytopenia. In one embodiment, the disorder is neutropenia. In another embodiment, the disorder is anemia. In some embodiments, the subject method comprises the step of detecting that a therapeutic benefit has been achieved. The ordinarily skilled artisan will appreciate that such measures of therapeutic efficacy will be applicable to the particular disease being modified, and will recognize the appropriate detection methods to use to measure therapeutic efficacy.

[000223] In another embodiment, the present invention includes a method of treating a disease in a subject in need thereof comprising providing to the subject an effective amount of cells transduced with a gene delivery vector, *e.g.*, a viral vector, that expresses a therapeutic gene product in the cells. In particular embodiments, the cells are autologous to the subject. In certain embodiments, the cells are erythroid cells, *e.g.*, hematopoietic stem cells or committed hematopoietic erythroid progenitor cells. In some embodiments, the cell is a bone marrow cell, *e.g.*, a lineage depleted bone marrow cell. In particular embodiments, the method is used to treat FA, and the viral vector is a LV comprising an expression construct disclosed herein comprising a human PGK promoter operably linked to a FANCA gene cDNA or coding sequence, and a mutated wPRE disclosed herein. In particular embodiments, the cells are provided to the subject parenterally, *e.g.*, via intravenous injection.

[000224] In another embodiment, the present invention includes a method of treating FA in a subject in need thereof, comprising providing to the subject an effective amount of autologous CD34⁺ stem cells transduced with a LV vector that expresses a FANCA cDNA in the cells, wherein the LV vector comprises a human PGK promoter operably linked to the FANCA cDNA or coding sequence, and a mutated wPRE sequence disclosed herein. In particular embodiments, the cells are hematopoietic stem cells or committed hematopoietic erythroid progenitor cells, *e.g.*, bone marrow cells. In particular embodiments, the cells are provided to the subject parenterally, *e.g.*, via intravenous injection.

[000225] Expression of the transgene using the subject transgene is expected to be robust. Accordingly, in some instances, the expression of the transgene, *e.g.* as detected by

measuring levels of gene product, by measuring therapeutic efficacy, etc. may be observed two months or less after administration, e.g. 4, 3 or 2 weeks or less after administration, for example, 1 week after administration of the subject composition. Expression of the transgene is also expected to persist over time. Accordingly, in some instances, the expression of the transgene, e.g. as detected by measuring levels of gene product, by measuring therapeutic efficacy, etc., may be observed 2 months or more after administration of the subject composition, e.g., 4, 6, 8, or 10 months or more, in some instances 1 year or more, for example 2, 3, 4, or 5 years, in certain instances, more than 5 years.

[000226] In certain embodiments, the method comprises the step of detecting expression of the transgene in the cells or in the subject, wherein expression is enhanced relative to expression from a polynucleotide cassette not comprising the one or more improved elements of the present disclosure. Typically, expression will be enhanced 2-fold or more relative to the expression from a reference, i.e. a control polynucleotide cassette, e.g. as known in the art, for example 3-fold, 4-fold, or 5-fold or more, in some instances 10-fold, 20-fold or 50-fold or more, e.g. 100-fold, as evidenced by, e.g. earlier detection, higher levels of gene product, a stronger functional impact on the cells, etc.

[000227] Typically, if the subject composition is an LV comprising the subject a polynucleotide cassette of the present disclosure, an effective amount to achieve a change in will be about 1×10^8 vector genomes or more, in some cases 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , or 1×10^{13} vector genomes or more, in certain instances, 1×10^{14} vector genomes or more, and usually no more than 1×10^{15} vector genomes. In some cases, the amount of vector genomes that is delivered is at most about 1×10^{15} vector genomes, e.g. 1×10^{14} vector genomes or less, for example 1×10^{13} , 1×10^{12} , 1×10^{11} , 1×10^{10} , or 1×10^9 vector genomes or less, in certain instances 1×10^8 vector genomes, and typically no less than 1×10^8 vector genomes. In some cases, the amount of vector genomes that is delivered is 1×10^{10} to 1×10^{11} vector genomes. In some cases, the amount of vector genomes that is delivered is 1×10^{10} to 3×10^{12} vector genomes. In some cases, the amount of vector genomes that is delivered is 1×10^9 to 3×10^{13} vector genomes. In some cases, the amount of vector genomes that is delivered is 1×10^8 to 3×10^{14} vector genomes.

[000228] In some cases, the amount of pharmaceutical composition to be administered may be measured using multiplicity of infection (MOI). In some cases, MOI may refer to the ratio, or multiple of vector or viral genomes to the cells to which the nucleic may be

delivered. In some cases, the MOI may be 1×10^6 . In some cases, the MOI may be 1×10^5 - 1×10^7 . In some cases, the MOI may be 1×10^4 - 1×10^8 . In some cases, recombinant viruses of the disclosure are at least about 1×10^1 , 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , 1×10^{13} , 1×10^{14} , 1×10^{15} , 1×10^{16} , 1×10^{17} , and 1×10^{18} MOI. In some cases, recombinant viruses of this disclosure are 1×10^8 to 3×10^{14} MOI. In some cases, recombinant viruses of the disclosure are at most about 1×10^1 , 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , 1×10^{13} , 1×10^{14} , 1×10^{15} , 1×10^{16} , 1×10^{17} , and 1×10^{18} MOI.

[000229] In some aspects, the amount of pharmaceutical composition comprises about 1×10^8 to about 1×10^{15} particles of recombinant viruses, about 1×10^9 to about 1×10^{14} particles of recombinant viruses, about 1×10^{10} to about 1×10^{13} particles of recombinant viruses, or about 1×10^{11} to about 3×10^{12} particles of recombinant viruses.

[000230] Any total number of viral particles suitable to provide appropriate transduction of cells to confer the desired effect or treat the disease can be administered to the mammal. In various preferred embodiments, at least 10^8 ; 5×10^8 ; 10^9 ; 5×10^9 ; 10^{10} ; 5×10^{10} ; 10^{11} ; 5×10^{11} ; 10^{12} ; 5×10^{12} ; 10^{13} ; 1.5×10^{13} ; 3×10^{13} ; 5×10^{13} ; 7.5×10^{13} ; 9×10^{13} , 1×10^{14} viral particles, or 5×10^{14} viral particles or more, but typically not more than 1×10^{15} viral particles are injected. Any suitable number of administrations of the vector to the mammal or the primate eye can be made. In one embodiment, the methods comprise a single administration; in other embodiments, multiple administrations are made over time as deemed appropriate by an attending clinician. In some embodiments at least 2×10^8 VG/ml of 5×10^5 cells/ml is required in a single administration (24 hours transduction) to result in high transduction efficiencies. Individual doses are typically not less than an amount required to produce a measurable effect on the subject, and may be determined based on the pharmacokinetics and pharmacology for absorption, distribution, metabolism, and excretion (“ADME”) of the subject composition or its by-products, and thus based on the disposition of the composition within the subject. This includes consideration of the route of administration as well as dosage amount. Effective amounts of dose and/or dose regimen can readily be determined empirically from preclinical assays, from safety and escalation and dose range trials, individual clinician-patient relationships, as well as in vitro and in vivo assays such as those described herein and illustrated in the Examples.

[000231] In some embodiments, the dose of cells patients receive by infusion will be that which is obtained from the transduction process. In various preferred embodiments, at least at least about 1×10^1 , 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 or more CD34⁺ cells/KG of patient weight are infused into the patient. In some embodiments, between 1×10^6 and 4×10^6 CD34⁺ cells/KG of patient weight are infused into the patient. In other embodiments, 3×10^5 and 4×10^6 CD34⁺ cells/Kg of patient weight are infused into the patient. In some embodiments, cells will be infused into the patient a single dose. In other embodiments, cells will be infused into the patient in multiple doses. Transduced cells may be infused immediately after the transduction process is completed.

[000232] Once integrated, the therapeutic protein (*e.g.*, human FANCA protein) is expressed by the cells. Transduced FA cells are genetically corrected, and thus able to activate the FA pathway by the mono-ubiquitination of FANCD2 and FANCI. These proteins migrate to areas of DNA damage, and in cooperation with other DNA repair proteins, promote the repair of the DNA in these cells, as occurs in healthy cells

[000233] As described in further detail in the Examples, preclinical *in vitro* data with BM samples from human FA patients has already shown the efficacy of an FANCA LV to correct the phenotype of these cells.

[000234] Accordingly, the present invention provides methods for treatment of the hematological manifestations of FA. In one embodiment, the hematological manifestation of FA is selected from one or more of BMF, thrombocytopenia, leukopenia, pancytopenia, neutropenia, and anemia. In a particular embodiment, the hematological manifestation is bone marrow failure (BMF), which appears in pediatric ages in most FA patients. In one embodiment, the hematological manifestation is thrombocytopenia. In another embodiment, the hematological manifestation is leukopenia. In one embodiment, the hematological manifestation is pancytopenia. In one embodiment, the hematological manifestation is neutropenia. In another embodiment, the hematological manifestation is anemia. In one embodiment, the hematological manifestation is a combination of two or more of BMF, thrombocytopenia, leukopenia, pancytopenia, neutropenia, and anemia.

[000235] An FANCA LV does not directly treat solid tumors that may be generated in more advanced stages of the disease. Nevertheless, the improvement of the hematological status of FA patients treated by hematopoietic gene therapy may also improve the immunological surveillance against the development of solid tumors. Therefore, an indirect

antitumor effect may also be generated as a consequence of the treatment of FA patients with an FANCA LV.

[000236] In order to achieve successful gene therapy in FA, it is beneficial to collect from a subject a "sufficient" number of hematopoietic stem cells (HSC).

[000237] In one embodiment, HSCs are obtained, or collected, from a bone marrow sample. In one embodiment, the bone marrow sample is depleted of erythrocytes. In some embodiments, the bone marrow sample is depleted of CD16⁺ white blood cells. In some embodiments, the cells remaining after depletion techniques are washed. In another embodiment, non-specific IgG is added to the washed cells. In some embodiments, the non-specific IgG is flebogamma. Subsequently, CD34⁺ cells may be selected from the washed cells. In one embodiment, CD34⁺ cells are selected from the bone marrow sample.

Selection methods for CD34⁺ cells may be positive selection, negative selection, or a combination thereof.

[000238] In another embodiment, HSCs are obtained from peripheral blood. In one embodiment, the peripheral blood sample is depleted of erythrocytes. In some embodiments, the blood sample is depleted of CD16⁺ white blood cells. In some embodiments, the blood cells remaining after depletion techniques are washed. In another embodiment, non-specific IgG is added to the washed cells. In some embodiments, the non-specific IgG is flebogamma. Subsequently, CD34⁺ cells may be selected from the washed cells. In one embodiment, CD34⁺ cells are selected from the peripheral blood sample. Selection methods for CD34⁺ cells may be positive selection, negative selection, or a combination thereof.

[000239] In some embodiments of the present invention, the HSCs are obtained from a subject following mobilization. Mobilization may be achieved by treating the subject with drugs or compounds that cause the movement of stem cells from the bone marrow into the blood. The stem cells can be collected and stored. In some embodiments, mobilization is achieved by treating the subject with G-CSF (filgrastin). In other embodiments, mobilization is achieved by treating the subject with plerixafor. In yet other embodiments, mobilization is achieved by treating the subject with a combination of filgrastin and plerixafor. (**Figure 11** and **Figure 14**)

[000240] In one embodiment, at least 1 to 4 x 10⁶ CD34⁺ corrected cells (*e.g.*, FANCA transduced HSCs) per kilogram of patient weight are administered to restore haematopoiesis in a non-conditioned FA patient. In some embodiments, the transduced cells are infused or

administered into the patient immediately after transduction. **(Figure 11)** In other embodiments, the transduced cells are frozen prior to infusing or administering into the patient. **(Figure 11)**

[000241] The genetic correction of HSCs from FA patients, followed by the autologous transplantation of these cells (hematopoietic gene therapy), is a good alternative for FA patients, particularly those lacking an HLA-identical sibling. In one embodiment, hematopoietic gene therapy is the preferred treatment regimen for a patient lacking an HLA-identical sibling. In another embodiment, hematopoietic gene therapy is the preferred treatment regimen for a patient that has an HLA-identical sibling.

[000242] All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. It is understood that the present disclosure supersedes any disclosure of an incorporated publication to the extent there is a contradiction.

[000243] It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely", "only" and the like in connection with the recitation of claim elements, or the use of a "negative" limitation.

[000244] The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

[000245] Because FA-A is the most frequent complementation group in FA patients (Casado *et al.*, 2007, Taniguchi *et al.*, 2006), vectors expressing the FANCA gene and/or the EGFP marker gene are the focus of the Examples; however, other FANCA genes may be utilized to similarly treat other complementation groups.

Examples

Example 1

FANCA lentiviral vectors.

[000246] Because of the safer integration pattern of LVs compared to RVs (Gonzalez-Murillo *et al.*, 2008; Modlich *et al.*, 2009; Montini *et al.*, 2006; Mitchell RS, Beitzel BF, Schroder AR *et al.* Plos Biol. 2004;2:E234; Montini E *et al.* J Clin Invest. 2009;119:964-975; Schroder AR *et al.* Cell. 2002;110:521-529), it was an aim to develop LVs as

therapeutic vectors to correct the phenotype of FA cells (Gonzalez-Murillo *et al.*, 2009). Additionally, since recent studies have shown that LVs harboring potent internal promoters can also trans-activate neighboring genes (Modlich *et al.*, 2009), LVs were considered for use in the clinic as a compromise between therapeutic efficacy and risks to trans-activate neighboring genes. The objective was, therefore, to define threshold levels of FANCA expression that could be therapeutic, in order to limit risks of gene trans-activation by the enhancer/promoter driving the expression of the therapeutic gene. The efficacy of the FANCA LV was first verified *in vitro* in FA-A LCLs, thereafter in primary BM samples from FA-A patients, and finally *in vivo* in a mouse model of FA-A.

[000247] To this aim, LVs expressing FANCA under the control of different promoters: vav, PGK, CMV and SFFV promoters, were constructed. **Figure 1** is a schematic of the medicinal product. **Figure 2A** shows a schematic representation of LVs expressing FANCA under the control of different internal promoters. Additionally we also investigated the influence of post-transcriptional WPRE elements, both in the expression level of *FANCA* and the therapeutic efficacy of the LVs. Initially, all LVs were packaged with the chimeric GALV-TR envelope. Titers of $1-2 \times 10^6$ tu/ml were routinely obtained, and transductions conducted at estimated MOIs of 1-2 tu/cell. **Figure 2B** shows a Western blot analysis of FANCA in FA-A cells transduced.

[000248] In order to determine the level of FANCA mRNA that was conferred by each vector, transduced FA-A lymphoblast cell lines (LCL) were selected with 30 nM MMC for 5 days. After the selection process, transduced FA-A LCLs contained 0.81 to 3.04 copies of the respective LV per cell (**Table 1**). Unselected FA-A LCLs transduced with EGFP-LVs and LCLs from a healthy donor (HD) were used as controls.

[000249] Total FANCA mRNA levels, as well as relative FANCA mRNA levels per LV copy number were determined in LCLs transduced with the different LVs (**Table 1**). Compared to FANCA mRNA levels observed in HD LCLs, similar levels of FANCA mRNA/copy were observed in FA-A cells transduced with vav-FANCA and PGK-FANCA LVs. CMV-FANCA and more significantly SFFV-FANCA LVs conferred supra-physiological levels of FANCA mRNA/copy (3.6 and 5.6 fold, respectively). PGK-FANCA LVs harboring the WPRE or the mutated WPRE* sequences (Schambach *et al.*, 2006) increased FANCA mRNA levels 2.3-2.6 fold compared to PGK-LVs without WPRE. Consistent with other studies (Schambach *et al.*, 2006; Zanta-Boussif MA, Charrier S, Brice-Ouzet A *Et al.*. Validation of a Mutated Pre

Sequence Allowing High and Sustained Transgene Expression While Abrogating WHV-X Protein Synthesis: Application to the Gene Therapy of WAS. *Gene Ther.* 2009;16:605-619; Zufferey R, Donello JE, Trono D, Hope TJ. Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element Enhances Expression of Transgenes Delivered by Retroviral Vectors. *J Virol.* 1999;73:2886-2892), the insertion of the WPRE or WPRE* sequences significantly increased FANCA mRNA levels in cells transduced with PGK-FANCA LVs. Since the wild type WPRE element encodes for a C-terminal truncated version of the hepatitis virus X protein which could mediate a tumorigenic effect (Bouchard MJ, Schneider RJ. The Enigmatic X Gene Of Hepatitis B Virus. *J Virol.* 2004;78:12725-12734.2004; Kingsman SM, Mitrophanous K, Olsen JC. Potential Oncogene Activity of the Woodchuck Hepatitis Post-Transcriptional Regulatory Element (WPRE). *Gene Ther.* 2005;12:3-4), the LV with the mutated WPRE* sequence (Schambach *et al.*, 2006), which lacks any residual open reading frame is considered more adequate for clinical uses.

[000250] **Table 1** FANCA expression levels in FA-A LCLS transduced with FANCA-expressing lentiviral vectors.

Cell Donor	Lentiviral Vector	FANCA copies/ cell	mRNA		Protein	
			Total FANCA	FANCA /Copy	Total FANCA	FANCA /Copy
HD	-	2*	1	0.5	1	0.5
FA-A	SFFV-EGFP	0	0.27 ± 0.03	-	-	-
FA-A	VAV-FANCA-WPRE	1.1 ± 0.33	0.51 ± 0.01	0.57 ± 0.20	0.75	0.59
FA-A	PGK-FANCA	3.04 ± 1.36	1.92 ± 1.32	0.54 ± 0.19	1.3 ± 0	0.53 ± 0.24
FA-A	PGK-FANCA-WPRE	3.01 ± 0.30	3.91 ± 2.19	1.24 ± 0.60	1.74 ± 0.26	0.57 ± 0.03
FA-A	PGK-FANCA-WPRE *	1.52 ± 0.05	2.20 ± 0.70	1.43 ± 0.41	1.7	0.54
FA-A	CMV-FANCA-EGFP	1.4 ± 0.28	2.42 ± 0.06	1.9 ± 0.41	0.96	0.58
FA-A	SFFV-FANCA-EGFP	0.81 ± 0.18	2.13 ± 0.26	2.78 ± 0.36	1.6	1.7

*To estimate levels of FANCA mRNA and FANCA protein per copy of FANCA, 2 copies of genomic FANCA were considered in healthy donor LCLs.

[000251] To test whether differences in FANCA mRNA expression were confirmed at the protein level, Western blot analyses (**Figure 2B**) were conducted with samples shown in **Table 1**. As it was done with FANCA mRNA determinations, FANCA protein values were related not only to protein loadings, but also to the provirus copy number determined in each transduced FA-LCL. Compared to FANCA levels/copy determined in HD-LCLs, essentially normal levels of FANCA/copy were conferred by all tested FANCA-LVs, except by the SFFV-FANCA LV. In this case, relative FANCA levels/copy were 3.4 fold higher to levels determined in HD-LCLs. Western blots re-stained with anti-FANCD2 showed that, while FA-A LCLs transduced with the control vector were not able to mono-ubiquitinate FANCD2, FA-A LCLs transduced with either type of FANCA-LVs expressed both the non-ubiquitinated as well as the mono-ubiquitinated forms of FANCD2, consistent with a functional FA pathway in these cells.

[000252] When FANCA expression levels conferred by the different LVs in FA-A LCLs were compared with those observed in healthy donor (HD) LCLs (with two copies of FANCA), we concluded that the insertion of two LV copies per cell may result in physiological levels of the therapeutic protein, except in the case of SFFV-LVs, which would confer supra-physiological levels of FANCA. Achieving this copy number may fit the requirements of a clinical trial of FA patients, where transduction efficacies of at least 50% may be desired because of the low number of progenitors present in the BM of these patients (Gonzalez-Murillo *et al.*, 2009, Jacome *et al.*, 2006, Kelly *et al.*, 2007; Larghero J, Marolleau JP, Soulier J *et al.* Hematopoietic Progenitor Cell Harvest and Functionality in Fanconi Anemia Patients. Blood. 2002;100:3051).

[000253] In order to analyze possible differences in the therapeutic efficacy of the different FANCA-expressing LVs, the efficiency of each vector to correct the MMC-hypersensitivity of FA-A lymphoblast cell lines (LCLs) was determined. To this aim, FA-A LCLs were transduced with the different LVs (**Figure 3A**) and then exposed to increasing concentrations of MMC. Thereafter, the viability of transduced cells was determined. As shown in **Figure 3A**, all tested FANCA-LVs were equally efficient to revert the hypersensitivity of FA-A LCLs. Similarly, all vectors promoted the generation of nuclear FANCD2 foci in MMC-treated cells (**Figure 3B**), consistent with a functional FA pathway in FA-A cells transduced with either FANCA-LV.

[000254] Example 2 Mouse model studies

[000255] To evaluate the repopulating properties of BM samples from FA patients, either genetically corrected or not, several groups have transplanted BM cells from FA patients into immuno-deficient mice. Nevertheless, in no instance have significant engraftments been reported, most probably because of the reduced number of hematopoietic progenitors and HSCs present in the BM of these patients.

[000256] To evaluate the *in vivo* effects of the medicinal product, a mouse model for FA-A, which contains deletion in the Fanca gene, was used. In contrast to FA patients, these animals do not develop evident hematological defects. Nevertheless, their BM progenitor cells are highly sensitive to MMC (Rio *et al.*, 2002), as it is also the case in FA patients. Therefore, to determine whether the FANCA LV (**Figure 4A**) stably corrected the phenotype of the HSCs from FA-A mice *in vivo*, BM cells from FA-A mice were transduced with FANCA LV and then transplanted into irradiated FA-A recipients (**Figure 4B**). To evaluate whether after the transplantation of genetically-corrected cells the phenotype of the hematopoietic progenitors was corrected, BM samples from transplanted FA-A mice were cultured in methylcellulose in the absence and the presence of MMC (**Figure 4C**).

Additional data confirming the integration of the therapeutic cassette is presented in **Figures 42-44**. Linear Amplification Mediated (LAM) PCR is a method to retrieve the integration sites of different integrating vectors into the genome. A PCR product that starts from the known sequence of the vector and extends through the unknown flanking genome is generated and sequenced to identify the position within the genome of the vector integration. **Figure 42** represents the LAM-PCR analysis of FANCA-LV insertion sites in FA hematopoietic stem cells (HSC). **Figure 43** depicts LAM-PCR results for tracking of FANCA-LV treated cells. **Figure 44** shows the clonal diversity of Fanca ^{-/-} recipients transplanted with LV-corrected HSCs.

[000257] After 1 month post-transplantation normal hematological counts were observed in these animals, showing no evident toxicity of the LV. At this time, the number of proviral FANCA copies per cell varied between 0.5 to 10 copies/cell (similar results were observed at 3 and 6 months post-transplantation) (**Figure 5**). To evaluate whether after the transplantation of genetically-corrected cells the phenotype of the hematopoietic progenitors was corrected, BM samples from transplanted FA-A mice were cultured in methylcellulose in the absence and the presence of MMC.

[000258] As shown in **Figure 6**, a significant correction of the MMC-hypersensitivity was observed in the hematopoietic progenitors from FA mice that were subjected to gene therapy with FANCA LV as compared to the SF1-EGFP LV control. FA-A bone marrow (BM) cells were transduced with PGK_FANCA-WPRE* or control SF1-EGFP LVs and transplanted into irradiated FA-A mice. At 7 months post-transplantation, BM samples were harvested and cultured in methylcellulose in the presence of increasing concentrations of mitomycin C (MMC).

[000259] These data therefore indicate that FANCA LV can revert the phenotype of hematopoietic progenitors from FA-A mice after *in vivo* transplantation. In safety terms, none of the 30 mice that were transplanted with BM cells previously transduced with FANCA LV have developed symptoms of myeloproliferative disorders or leukemia (data obtained up to 1 year post-transplantation).

[000260] Example 3. Efficient transduction of fresh hematopoietic progenitors from FA patients with lentiviral vectors.

[000261] Because transduction of FA cryopreserved hematopoietic stem cells (HSC) was previously shown to be less efficient compared to transduction of fresh grafts (Jacome *et al.*, 2009), an effort to optimize the efficacy of the transduction of cryopreserved FA BM samples was undertaken. As shown in **Figure 7**, conducting three transduction cycles significantly increased the transduction efficacy of FA CFCs compared to values obtained after a single transduction cycle ($45.7 \pm 4.2\%$ versus $13.5 \pm 5.1\%$, respectively). Samples were subjected to standard transductions consisting in a single transduction cycle (16h) after 2h of static preloading (white bars; 1xS) or improved transduction consisting in three transduction cycles (2h+2h+12h) with the lentiviral vectors (grey bars; 3xD).

[000262] Example 4. PGK-FANCA-WPRE* lentiviral vector efficiently corrects the phenotype of bone marrow progenitors from FA-A patients.

[000263] Because of the efficacy of LVs in which FANCA was driven by the PGK promoter, and based on previous studies showing the stability (Follenzi A *et al.* Nat Genet.2000;25:217-222) and low genotoxic properties of LVs carrying the PGK promoter (Modlich *et al.*, 2009, Montini *et al.*, 2006, Montini *et al.*, 2009) further experiments were conducted in LCLs and also in bone marrow cells from FA-A patients using PGK-FANCA

LVs, free from any WPRE element or harboring the WPRE or WPRE* sequences. As shown in **Figure 8A**, all these three vectors conferred the same reversion in the hypersensitivity of FA-A LCLs to MMC.

[000264] To compare the efficacy of SFFV-FANCA and PGK-FANCA LVs to correct the phenotype of hematopoietic progenitors from FA-A patients, erythrocyte-depleted BM samples from FA-A patients were transduced as recently described (Jacome *et al.*, 2009). Erythrocyte-depleted BM cells were transduced for 16h in plates preloaded with LV supernatants as recently described (Gonzalez-Murillo *et al.*, 2009). Fourteen days later, the number of colonies grown in the absence and the presence of 10 nM MMC was scored to determine the proportion of progenitors that became resistant to the drug.

[000265] As shown in **Figure 8B**, when samples were transduced with EGFP-LVs almost no colonies were generated in the presence of MMC. In contrast to this observation, the transduction of FA-A BM cells with SFFV-FANCA and PGK-FANCA LVs allowed the growth of 26 and 38% of the colonies scored in cultures without MMC. The efficacy of PGK-FANCA LVs harboring the WPRE and WPRE* sequences was compared to WPRE-free PGK-FANCA LVs. As shown in **Figure 8B**, all the three LVs mediated a high and similar level of protection to MMC. Although the insertion of the post-transcriptional regulatory element WPRE* (Schambach *et al.*, 2006) was not necessary for improving the efficacy of PGK-FANCA LVs, this element will offer a redundant element to maintain in the long-term therapeutic levels of ectopic FANCA in the patient.

[000266] Taken together, data obtained in these studies strongly suggest that PGK-FANCA-WPRE* LVs confer sufficient levels of FANCA expression to correct the hematopoietic phenotype of FA-A patients. These results, together with previous observations showing the stability (Follenzi *et al.*, 2000) and safety properties of PGK-LVs (Modlich *et al.*, 2009, Montini *et al.*, 2006, Montini *et al.*, 2009) and the efficacy of the mutated WPRE* post-transcriptional element (Schambach *et al.*, 2006), reinforces the hypothesis that the PGK-FANCA-WPRE* LV may constitute an efficient and safe vector for the gene therapy of FA-A patients.

[000267] Example 5. Analysis of the efficacy of the GALV-TR and VSV-G packaged lentiviral vectors to transduce hematopoietic progenitors from FA patients.

[000268] Because LVs can be pseudotyped both with GALV-TR and VSV-G envelopes, new experiments were conducted in which unselected bone marrow cells from FA patients were transduced under optimized conditions with EGFP-LVs pseudotyped in the two envelopes.

[000269] For GALV-TR pseudotyped LVs, two transduction cycles with nonconcentrated LVs (estimated titer: 2×10^5 IUs/mL; estimated MOI: 2 IUs/cell) were conducted.

[000270] For VSV-G pseudotyped LVs, one transduction cycle with concentrated and purified LVs (estimated titer: 10^8 IUs/mL; estimated MOI: 50 IUs/cell).

[000271] As shown in **Figure 9**, under both conditions similar transductions of hematopoietic progenitors from FA patients were obtained, indicating that manufacturing compromises can determine the best envelope for packaging.

Example 6. Safety studies.

[000272] The transforming potential of the lentiviral vectors shown in **Figure 10A** was measured in replating frequency over copy number. As shown in **Figure 10B**, the transforming potential of the lentiviral backbone corresponding to the PGK-FANCA-WPRE* LV (a PGK-derived lentiviral vector) is markedly lower compared to the transformation capacity of vectors harboring viral promoters, which are the ones already used in the X1-SCID and CGD clinical trials.

[000273] With respect to *in vivo* studies with mice transplanted with BM cells transduced with the PGK-FANCA-WPRE* LV, 30 mice have been transplanted, and so far (up to 1 year post-transplantation) none of the transplanted animals have developed symptoms of myeloproliferative disorders or leukemia.

[000274] Example 7. Pharmaceutical product

[000275] The FANCA lentiviral vector is a third generation self-inactivated rHIV1- derived vector encoding the human FANCA cDNA under control of the human PGK promoter and regulated at the post-transcriptional level by a mutated WPRE lacking the X protein ORF (See **Figure 1**, **Figure 41**, and SEQ ID NO: 24). Such a FANCA lentiviral vector presents several advantages over the gamma-retroviral vectors previously used in FA gene therapy, notably the ability to transduce cells in spite of short pre-activation protocols which is advantageous to preserve the multi-lineage potential of hematopoietic stem cells.

[000276] For the generation of the therapeutic lentiviral vector, 293T cells transfected with three additional plasmids, providing all the required helper proteins for the packaging of the vector (See **Figures 38-40**) will comprise the final packaging of the PGK-FANCA-WPRE* medicinal product (**Figure 41**).

[000277] The PGK-FANCA-WPRE*LV therapeutic cassette comprises the human PGK promoter, the coding sequence for FANCA cDNA, and the WPRE* enhancer and comprises nucleotides 3541 to 9178 of SEQ ID NO: 24. The region of the transfer cassette comprising the human CMV immediate early promoter, the HIV packaging sequence, the ga and RRE elements, the therapeutic cassette, and the HIV self inactivating 3'LTR wherein the therapeutic cassette comprises the human PGK promoter, the coding sequence for FANCA cDNA, and the WPRE* enhancer is coded for by nucleotides 1586-9495 of SEQ ID NO: 24.

[000278] Nucleotides 1586-1789 of SEQ ID NO: 24 comprise human CMV immediate early promoter. Nucleotides 2031-2156 of SEQ ID NO: 24 comprise HIV1 psi packaging signal. Nucleotides 2649-2882 of SEQ ID NO: 24 comprise HIV1 RRE element. Nucleotides 3378-3495 of SEQ ID NO: 24 comprise HIV cPPT/CTS element. Nucleotides 3541-4051 of SEQ ID NO: 24 comprise the hPGK promoter. Nucleotides 4078-8445 of SEQ ID NO: 24 comprise human FANCA-A cDNA. Nucleotides 8502-9178 of SEQ ID NO: 24 comprise mutated WPRE element. Nucleotides 9262-9495 of SEQ ID NO:24 comprise the HIV delta U 3' LTR.

[000279] Example 8. Clinical Study FANCOSTEM

[000280] In the U.S., two gene therapy trials have been conducted in FA-A and FA-C patients, which showed no clinical efficacy (Liu, J. M., et al. (1999). Engraftment of hematopoietic progenitor cells transduced with the Fanconi anaemia group C gene (FANCC). *Hum. Gene Ther.* 10: 2337-2346; Kelly, P. F., et al. (2007). Stem cell collection and gene transfer in fanconi anaemia. *Mol Ther* 15: 211-219). The efficacy of the aforementioned trials could be significantly improved through various optimizations. Two clinical trials were conducted in tandem to determine the feasibility of a process for collecting and purifying a sufficient number of CD34⁺ cells for future clinical use (FANCOSTEM) and, in parallel, to evaluate the safety and efficacy of the gene therapy in patients with FA complementation group A (FA-A) (FANCOLEN) See **Figure 11**.

[000281] The main inclusion criteria for FANCOSTEM were Patients with a diagnosis of AF, confirmed by a test of chromosomal instability with diepoxybutane or mitomycin C, Age > 1

year, and At least one of the following parameters should be as high as: 1) Hemoglobin: 8.0 g/dL, 2) Neutrophils: 750/mm³, 3) Platelets: 30,000/mm³. Ten patients were recruited, nine were screened out of which 2 failed (mosaic patients.) **Figure 12** shows the haematological parameters of recruited patients. Seven (7) patients were treated with G-CSF and Plerixafor. The mobilization regimen utilized the administration of G-CSF (neupogen; 12 µg/Kg/12 hours) and plerixafor (mozobil; 240 µg/kg body weight/day). Mobilized peripheral blood (mPB) CD34+ cells were transduced under GMP conditions with a PGK-FANCA-WPRE* LV using a short *ex vivo* transduction protocol (**Figure 13**). While two patients who were 15 and 16 years old did not reach the threshold level of CD34+ cells in peripheral blood, apheresis could be conducted in five patients with ages between 3-5 years old. In these patients the median number of CD34+ cells/kg was of 6.6x10⁶ (range: 1.6x10⁶ to 7.6x10⁶). After CD34+ cell selection, the number of CD34+ cells/kg was 2.0x10⁶ (range: 8.5x10⁵ and 5.1x10⁶). No severe adverse events related to the mobilization regimen have been detected in any treated patient followed for up to 2.5 years.

[000282] **Figure 14** shows G-CSF/Plerifaxor-mediated Mobilization of CD34+ cells in FA-A patients and **Figure 15** shows G-CSF/Plerifaxor-mediated Mobilization of colony forming cells (CFC) in FA-A patients. **Figure 16A** shows CD34+ cell collection in FANCOSTEM and **Figure 16B** shows compared to previous studies.

[000283] **Figure 17** shows comparison between predicted CD34+ cell numbers in bone marrow (BM) vs actual numbers in mobilized peripheral blood (mPB). **Figure 18** is a summary of the CD34+ cells collection in G-CSF/ Plerixafor mobilized FA-A patients. The number of CD34+ after selection correlates with the number of CD34+ cells/µl in BM at day 0 (data not shown).

[000284] **Figure 19** depicts CD34 expression prior to and after immunoselection of mPB CD34+ cells from healthy donor (HD) and FA patients.

[000285] From these data, we concluded that compared to other clinical studies, evident improvements in the collection of CD34+ cells have been observed so far in patients treated with Filgrastin (G-CSF) and Plerixafor and only FA patients in early stages of the disease seem to be suitable for the collection of clinically relevant numbers of HSCs.

[000286] Example 9. Clinical Study FANCOLEN

[000287] The second parallel clinical trial (FANCOLEN) aimed to evaluate the safety and efficacy of the gene therapy in patients with FA complementation group A (FA-A) (FANCOLEN). In order to restore the hematopoiesis of FA patients by the infusion of gene corrected autologous HSCs, optimized vectors and transduction protocols were developed. Specifically, this was a Phase I/II clinical trial to evaluate the safety and efficacy of the infusion of autologous CD34⁺ cells transduced with a lentiviral vector carrying the FANCA gene (Orphan drug) for patients with Fanconi Anemia Subtype A.

[000288] The main inclusion criteria were patients with a diagnosis of FA-A, Age > 1 year, and at least one of the following parameters should be below the threshold of: 1) Hemoglobin: 8.0 g/dL; 2) Neutrophils: 1,000/mm³; 3) Platelets: 50,000/mm³.

[000289] More specific subject inclusion criteria include 1) Patients of the complementation group FA-A; 2) At least one of the following parameters must be lower than the values indicated: haemoglobin: 8.0 g/dL; neutrophils: 750/mm³; platelets: 30,000/mm³; 3) Minimum age: 1 year; 4) Maximum age: 21 years; 5) Lansky index > 60% 6) Mild organ functional impairment; 7) Provide informed consent in accordance with current legislation; 8) Number of cells to transduce: at least 3x10⁵ purified CD34⁺ cells/Kg of patient weight; 9) Women of childbearing age must have a negative urine pregnancy test at the baseline visit, and accept the use of an effective contraception method during participation in the trial.

[000290] Subject exclusion criteria include 1) Patients with an HLA-identical family donor; 2) Evidence of myelodysplastic syndrome or leukemia, or cytogenetic abnormalities predictive of these conditions in bone marrow aspirate analysis. This assessment should be made by valid studies two months before the patient enters the clinical trial; 3) Evidence that the patient has signs of somatic mosaicism with improved haematology; 4) Any concomitant disease or condition that, in the opinion of the investigator, deems the subject unfit to participate in the study; 5) Pre-existing sensory or motor impairment > = grade 2 according to the criteria of the National Cancer Institute (NCI); 6) Pregnant or breastfeeding women.

[000291] Route of administration: Patients received the cells transduced with therapeutic vector by intravenous infusion.

[000292] Dose of cells: The dose of cells patients received by transfusion was that which was obtained from transduction, between 3x10⁵ and 4x10⁶ purified CD34⁺ cells/kg of patient weight.

[000293] Below 3x10⁵ CD34⁺ cells/Kg is highly unlikely to produce a patient graft from transduced cells, especially considering that this clinical trial will initially infuse unconditioned

patients. In the gene therapy trial in Fanconi anemia patients conducted by Dr Williams (Kelly *et al.*, 2007) the number of cells infused in 2 patients (FAAGT1001 and 1003) was 4.5×10^5 and 3.25×10^5 nucleated cells/kg of patient weight, respectively. In both patients, a transient improvement in Hb and platelet count was seen, but without being able to demonstrate an associated presence of the transgene. Unlike in Dr Williams' trial where cells were transduced for 4 days with a gamma-retroviral vector, in this clinical trial cells were transduced with a more efficacious lentiviral vector, and the transduction will be conducted for a maximum of 48 hours, *i.e.*, much shorter than the 4 days used in the previous protocol. Given this, we considered 3×10^5 purified CD34⁺ cells/kg of patient body weight to be a reasonable lower limit for this trial.

[000294] The upper limit of 4×10^6 purified CD34⁺ cells/kg is not based on the need to limit the number of cells infused, as a greater number of cells increase the likelihood of graft. Rather the limit of 4×10^6 purified CD34⁺ cells/kg of patient body weight comes from the difficulty in mobilizing and collecting cells exceeding this number from patients with Fanconi anemia, characterized by having a reduced CD34⁺ cell count in their bone marrow (Jacome *et al.*, 2009).

[000295] Dosage regimen: The cells transduced with the therapeutic vector were infused in a single dose to the patient. This was for two main reasons: 1) All hematopoietic gene therapy trials conducted so far have been performed using a single infusion of transduced cells (Naldini, 2011). Following this prior experience, we were not inclined to vary this parameter. 2) Single infusion of all transduced cells would increase the likelihood that there is a greater graft, compared to infusion of the same dose fractionated.

[000296] Recruitment period: Patients are recruited over a period of 3 years. As patients with subtype FA-A are the most frequent in the FA patient population (around 80% of Spanish patients with FA correspond to this complementation group (Casado *et al.* (2007))), the constructed therapeutic vector carries the *FANCA* gene. Therefore, of FA patients only those patients belonging to subtype FA-A can participate in this study. Any patients in this complementation group, paediatric or adult, provided they meet the defined inclusion and exclusion criteria, were included in the study.

[000297] Specific description of the primary and, if any, secondary variables that will be assessed in the clinical trial include 1) Main variable: A) To determine the toxicity associated with infusion of autologous CD34⁺ cells transduced with the therapeutic lentiviral vector in patients with Fanconi anemia subtype A. B) To determine the degree of grafting associated with the infusion of autologous CD34⁺ cells transduced with the therapeutic lentiviral vector in patients with Fanconi

anemia subtype A 2) Secondary variables: To determine the clinical response associated with infusion of autologous CD34⁺ cells transduced with the therapeutic lentiviral vector in patients with Fanconi anemia subtype A.

[000298] CD34⁺ cells from bone marrow and/or mobilized in peripheral blood (fresh and/or cryopreserved) from patients with Fanconi anemia subtype A (FA-A) were transduced *ex vivo* with a lentiviral vector carrying the *FANCA* gene (orphan drug) (**Figure 11**). After cell transduction, patients received an infusion of these genetically corrected stem cells in order to restore haematopoiesis.

[000299] Assessment: Patients were assessed before initiating treatment, obtaining prior informed consent. The assessment was carried out in the month before infusion of genetically corrected cells, through a standard physical examination (including weight and height), peripheral blood cell counts, basic biochemistry, and a bone marrow aspirate.

[000300] Transduction and infusion of genetically corrected CD34⁺ cells: The purified CD34⁺ cell population was transduced *ex vivo* with the therapeutic lentiviral vector. After cell transduction, product quality control evaluations was carried out, aliquots were cryopreserved for further study, and the product was prepared for infusion into patients.

[000301] If in two patients infused with an acceptable number of transduced cells (at least 1 million cells infused/kg of body weight, in an aliquot in which at least 0.3 copies of the vector/cell is detected after at least 7 days in culture *in vitro*) AT LEAST 0.1 COPIES VECTOR/CELL is not observed in either the bone marrow or peripheral blood at 6 MONTHS POST-INFUSION, the subsequent patients will be subject to conditioning process prior to cell infusion.

[000302] For a patient to be eligible for conditioning of any kind there must be a suitable method of rescuing potential aplasia of bone marrow associated with conditioning and possible implant failure of transduced cells. Rescue methods include a unit of umbilical cord blood or hematopoietic cells from a haploidentical donor. Cells will be infused intravenously.

[000303] The dose of cells patients receive by infusion was that which is obtained from the transduction process, between 3×10^5 and 4×10^6 CD34⁺ cells/Kg of patient weight.

[000304] Cells are infused into the patient a single dose.

[000305] Transduced cells are infused immediately after the transduction process is completed.

[000306] The product infused consists of a suspension of CD34⁺ cells which was packaged in a sterile bag for infusion by the CLINISTEM GMP laboratory at CIEMAT.

[000307] Recruitment Period: 3 years from infusion of the 1st patient.

[000308] Follow-up period: two years from the infusion of transduced cells. However, patients are monitored outside the clinical trial for a period of 10 years.

[000309] Monitoring of the graft of transduced cells will be carried out on peripheral blood and bone marrow samples.

[000310] First 72 hours after infusion: During this period, vital signs were recorded every 8 hours and vital organ functions were monitored (electrolyte profile, haematology, renal and hepatic function) every 24 hours.

[000311] Subsequently the following checks were carried out as shown in **Table 2**.

[000312] **Table 2.**

	<i>Peripheral blood</i>												
	Week:				Month:								
Haematological monitoring	2	4	6	-	2	4	6	9	12	15	18	21	24
Copies of the therapeutic vector	2	4	6	-	2	4	6	9	12	15	18	21	24
	<i>Bone marrow</i>												
	Week:			Month:									
Cytology	4		-		6			12					24
Copies of the therapeutic vector	4		-		6			12					24

[000313] Patients diagnosed with FA and belonging to complementation group FA-A will be included in the study. Patients were considered if cellular phenotype correction has been demonstrated by transduction with vectors carrying the *FANCA* gene or if bi-allelic pathogenic mutations in this gene are demonstrated.

[000314] Patient FA 02005 fit the criteria for FANCOSTEM and FANCOLEN as summarized in **Figure 20**. **Figure 21** depicts the tests showing FA diagnosis of patient FA-02005 prior to gene therapy and **Figure 22** shows the follow up of the cell manufacturing process for patient FA-02005. **Figure 23** shows vector copy number in patient FA02005 prior to gene therapy and 2 weeks, 4 weeks, 6 weeks, 2 months, 3 months, 4 months, and 5 months after gene therapy. Follow-up of the first not-conditioned FA-A patient prior to and after gene therapy, patient

FA02005 (4 year old) is represented for hemoglobin (**Figure 24**), neutrophils (**Figure 25**), and platelets (**Figure 26**).

[000315] For patient FA02002, the hematological evolution is presented in **Figure 27**, diagnosis is presented in **Figure 28**, and follow up of the cell manufacturing process is presented in **Figure 29**. **Figure 30** shows analysis of CD34 expression in healthy donor and FA mPB during the different steps required for LV-transduction for patient FA-02002. **Figure 31** shows vector copy number in patient FA02002 prior to gene therapy and 2 weeks, 4 weeks, 6 weeks after gene therapy. Follow-up of the patient FA02002 (infused with cryopreserved cells) is presented for hemoglobin (**Figure 32**), neutrophils (**Figure 33**), and platelets (**Figure 34**).

[000316] The data for these two patients support the conclusion that the patients were infused with a significant number of gene-corrected mobilized peripheral blood (mPB) CD34⁺ cells. No serious adverse events have been observed in any of the two treated patients. Gene marking levels of around 1-5 copies/1000 peripheral blood cells were detected in treated patients at 15 days to 5 months post-GT, with moderate increases along time. These viral copy numbers were around 100x higher compared to the highest value detected at 3 weeks post-GT in previous trials (3 copies/10⁵ cells; Kelly *et al.* 2007).

[000317] Example 6. Transduction of Fresh mPB CD34⁺ Cells from FA-A Patients

[000318] The short transduction with the therapeutic vector of small aliquots of mobilized peripheral blood (mPB) CD34⁺ samples from patients treated with the G-CSF/Plerixafor regimen showed transduction efficacies between 17-45% (**Figure 35**). Small aliquots of these samples were transplanted into NSG mice conditioned with 1.5 Gy. Most of the transplanted samples engrafted into the NSG mice (1-10% of the BM cells were hCD45⁺/mCD45⁻) (**Figure 36**). Moreover, an evident selection advantage of corrected CD34⁺ FA-A cells was observed in engrafted mice (**Figure 37**).

[000319] These results show that transduced FA-A mPB CD34⁺ cells engraft into NSG mice and there is an *in vivo* proliferation advantage of corrected human FA-A repopulating cells takes place in NSG recipient mice.

Throughout this specification, unless the context requires otherwise, the word "comprise" or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or method step or group of elements or integers or method steps but not the exclusion of any element or integer or method step or group of elements or integers or method steps.

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgement or admission or any form of suggestion that the prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

CLAIMS

1. A method of treating Fanconi anemia in a subject in need thereof, comprising administering a cell comprising a recombinant gene delivery vector, the recombinant gene delivery vector comprising an expression cassette comprising a polynucleotide sequence comprising in the following 5' to 3' order:

(a) a human phosphoglycerate kinase (PGK) promoter sequence or a functional homolog or variant thereof; and

(b) a sequence encoding a human FANCA polypeptide or a functional fragment or variant thereof;

wherein the sequence encoding the human FANCA polypeptide or functional fragment or variant thereof is operably linked to the PGK promoter sequence, and wherein the cell is autologous to the subject.

2. The method of claim 1, wherein the expression cassette further comprises a woodchuck hepatitis virus regulatory element (WPRE) RNA export signal sequence or a functional variant or fragment thereof.

3. The method of claim 1 or claim 2, wherein the FANCA polypeptide or functional fragment or variant thereof comprises the sequence set forth in SEQ ID NO: 25.

4. The method of any one of claims 1 to 3, wherein the sequence encoding the FANCA polypeptide or functional fragment or variant thereof comprises the sequence set forth in SEQ ID NO: 8.

5. The method of any one of claims 1 to 4, wherein the PGK promoter comprises a nucleotide sequence of SEQ ID NO: 7.

6. The method of any one of claims 2 to 5, wherein the WPRE element comprises a nucleotide sequence of SEQ ID NO: 23.

7. The method of any one of claims 1 to 6, wherein the cassette comprises the nucleotide sequence of SEQ ID NO: 24.

8. The method of any one of claims 1 to 7, wherein the cassette further comprises one or more enhancer sequences.
9. The method of any one of claims 1 to 8, wherein the cassette further comprises:
 - (d) a polypurine tract (PPT) or polyadenylation (polyA) signal sequence.
10. The method of any one of claims 1 to 9, wherein the cassette further comprises one or more of the following sequences:
 - (e) a packing signal sequence;
 - (f) a truncated Gag sequence;
 - (g) a Rev responsive element (RRE);
 - (h) a central polypurine tract (cPPT);
 - (i) a central terminal sequence (CTS); and
 - (j) an upstream sequence element (USE), optionally from simian virus 40 (SV40-USE).
11. The method of any one of claims 1 to 10, wherein the recombinant gene delivery vector is a virus or viral vector.
12. The method of claim 11, wherein the virus or viral vector is a lentivirus (LV).
13. The method of any one of claims 1 to 12, wherein the cell is a hematopoietic stem cell.
14. The method of claim 13, wherein the hematopoietic stem cell is a CD34⁺ cell.
15. The method of claim 14, wherein the CD34⁺ cell was obtained from the subject after the subject was treated with a combination of: (i) G-CSF or Filgrastim; and (ii) Plerifaxor.
16. The method of claim 14 or claim 15, wherein the CD34⁺ cell was transduced with the recombinant gene delivery vector comprising the expression cassette.

17. The method of any one of claims 14 to 16, wherein the CD34⁺ cell was transduced by contacting the CD34⁺ cell with the recombinant gene delivery vector for about 24 hours.

18. The method of any one of claims 1 to 17, wherein the method inhibits the development of, halts progression of, and/or reverses progression of a hematological manifestation of Fanconi anemia in the subject.

19. The method of claim 18, wherein the hematological manifestation of Fanconi anemia is selected from one or more of BMF, thrombocytopenia, leukopenia, pancytopenia, neutropenia, and anemia.

20. Use of CD34⁺ cells comprising an expression cassette in the manufacture of a medicament for treating Fanconi anemia in a subject in need thereof, wherein the expression cassette comprises a polynucleotide sequence comprising in the following 5' to 3' order:

(a) a human phosphoglycerate kinase (PGK) promoter sequence or a functional homolog or variant thereof; and

(b) a sequence encoding a human FANCA polypeptide or a functional fragment or variant thereof;

wherein the sequence encoding the human FANCA polypeptide or functional fragment or variant thereof is operably linked to the PGK promoter sequence, and wherein the cell is autologous to the subject.

21. A method for producing an autologous cell comprising a recombinant gene delivery vector for administration to a subject in need thereof, the method comprising:

- (i) obtaining a biological sample comprising CD34⁺ cells from the subject;
- (ii) preparing a cell population enriched for CD34⁺ cells from the biological sample;
- (iii) transducing the cell population enriched for CD34⁺ cells with the recombinant gene delivery vector;

wherein the recombinant gene delivery vector comprises an expression cassette comprising a polynucleotide sequence comprising in the following 5' to 3' order:

(a) a human phosphoglycerate kinase (PGK) promoter sequence or a functional homolog or variant thereof; and

(b) a sequence encoding a human FANCA polypeptide or a functional fragment or variant thereof;

and wherein the sequence encoding the human FANCA polypeptide or functional fragment or variant thereof is operably linked to the PGK promoter sequence,

22. The method of claim 21, wherein the transducing comprises contacting the cell population enriched for CD34⁺ cells with the vector for about 24 hours.

23. The method of claim 21 or claim 22, wherein the subject is provided a combination of: (i) G-CSF or Filgrastim; and (ii) Plerifaxor to mobilize CD34⁺ cells within the subject, prior to obtaining the biological sample.

24. The method of any one of claims 21 to 23, wherein preparing the cell population comprises depleting erythrocytes.

25. The method of any one of claims 21 to 24, wherein preparing the cell population comprises enriching for CD34⁺ cells by positive selection, negative selection, or a combination thereof.

26. An autologous cell produced according to the method of any one of claims 21 to 25.

27. A pharmaceutical composition comprising the cell of claim 26.

Figure 1.

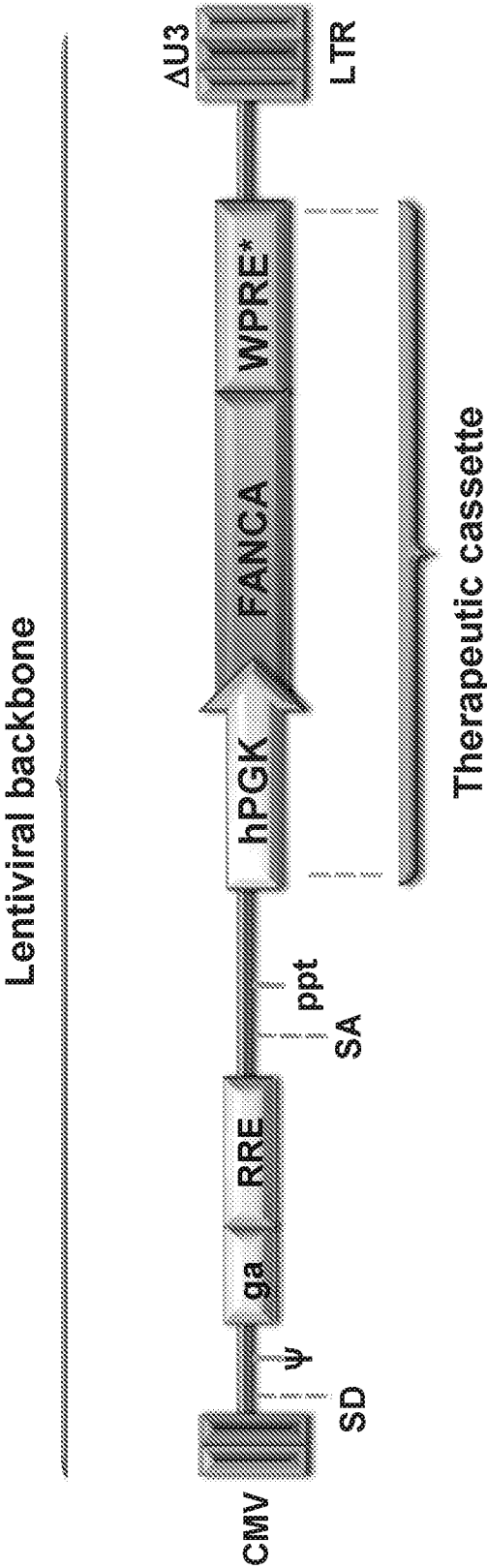


Figure 2.

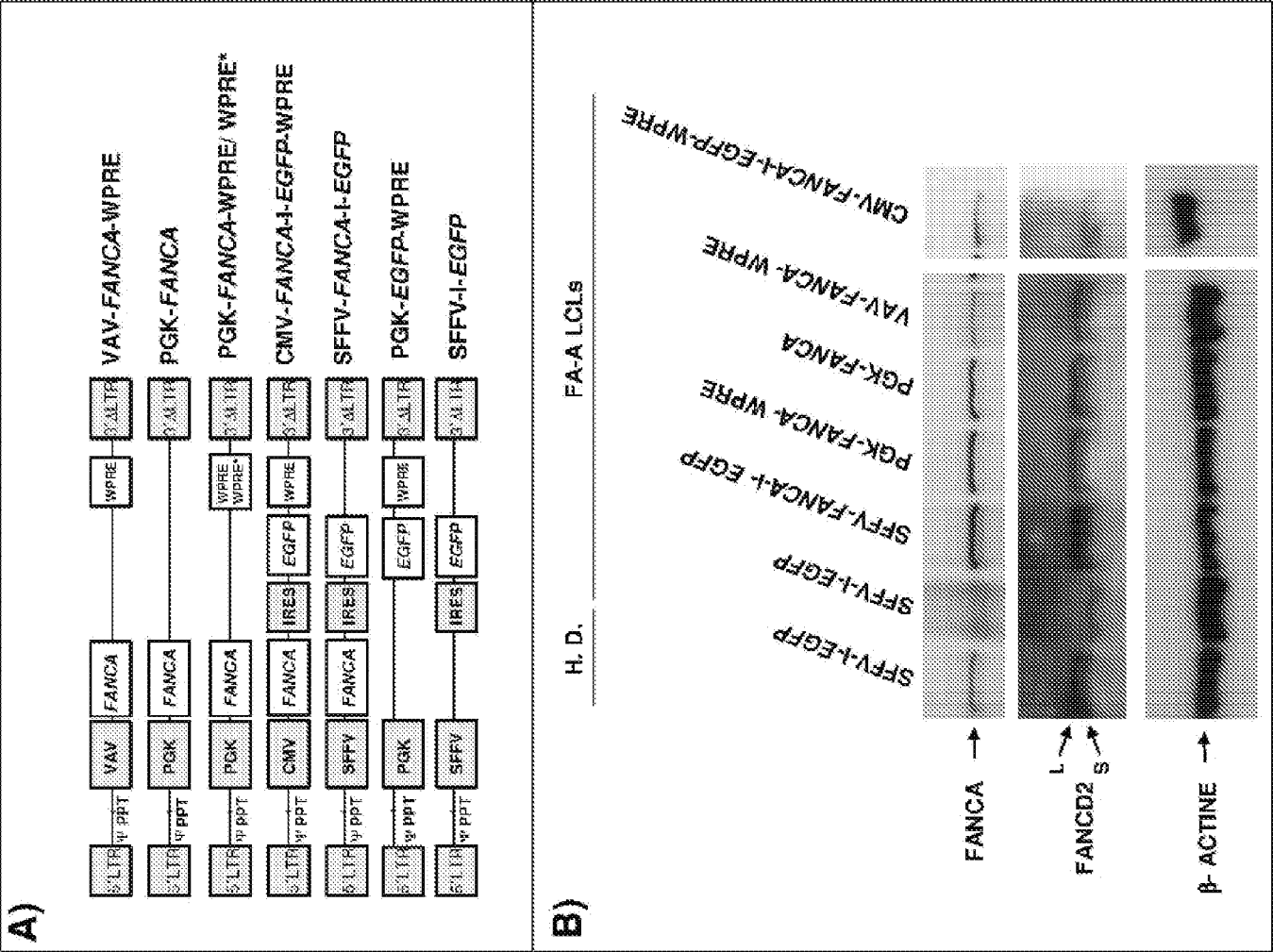


Figure 3.

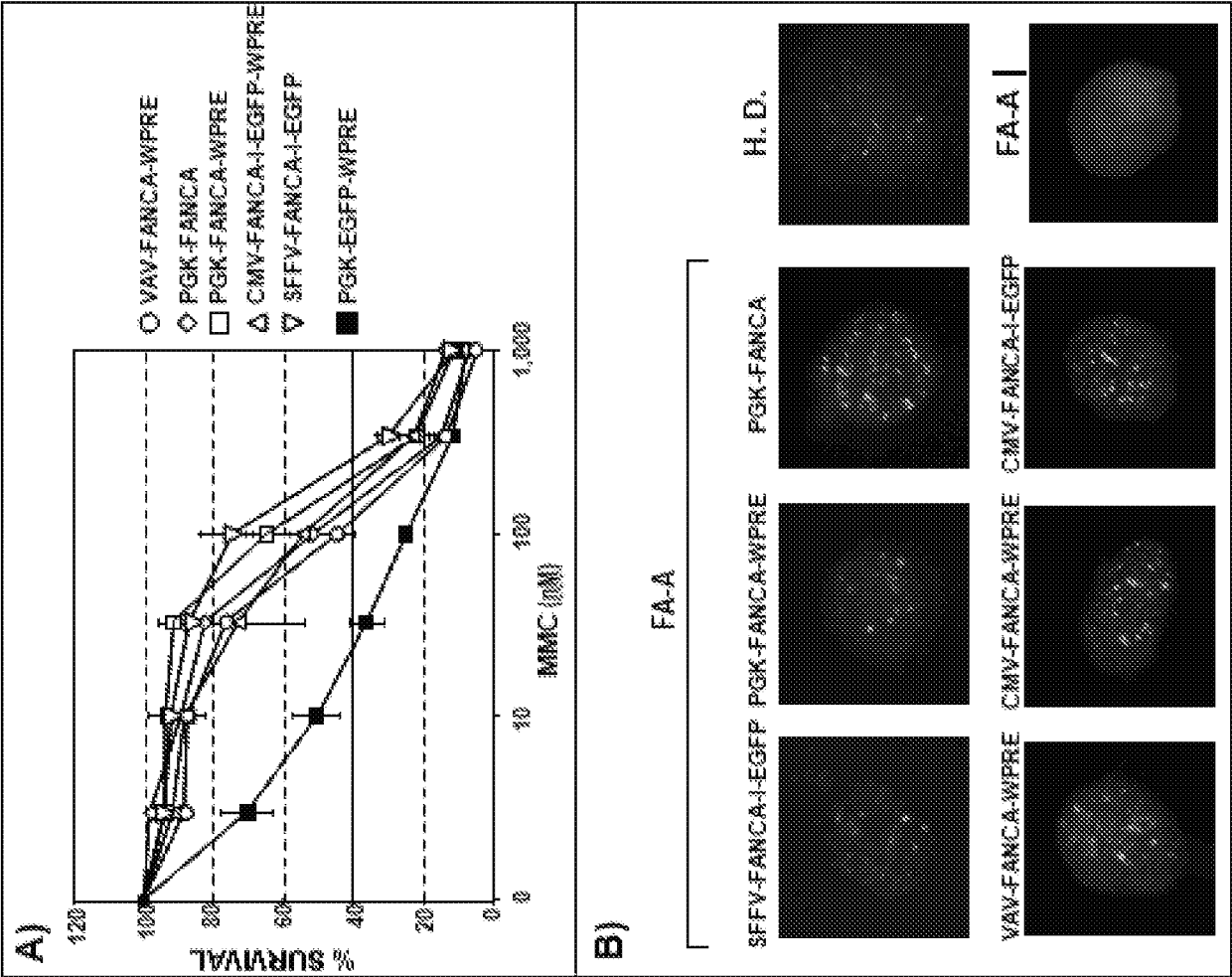


Figure 4.

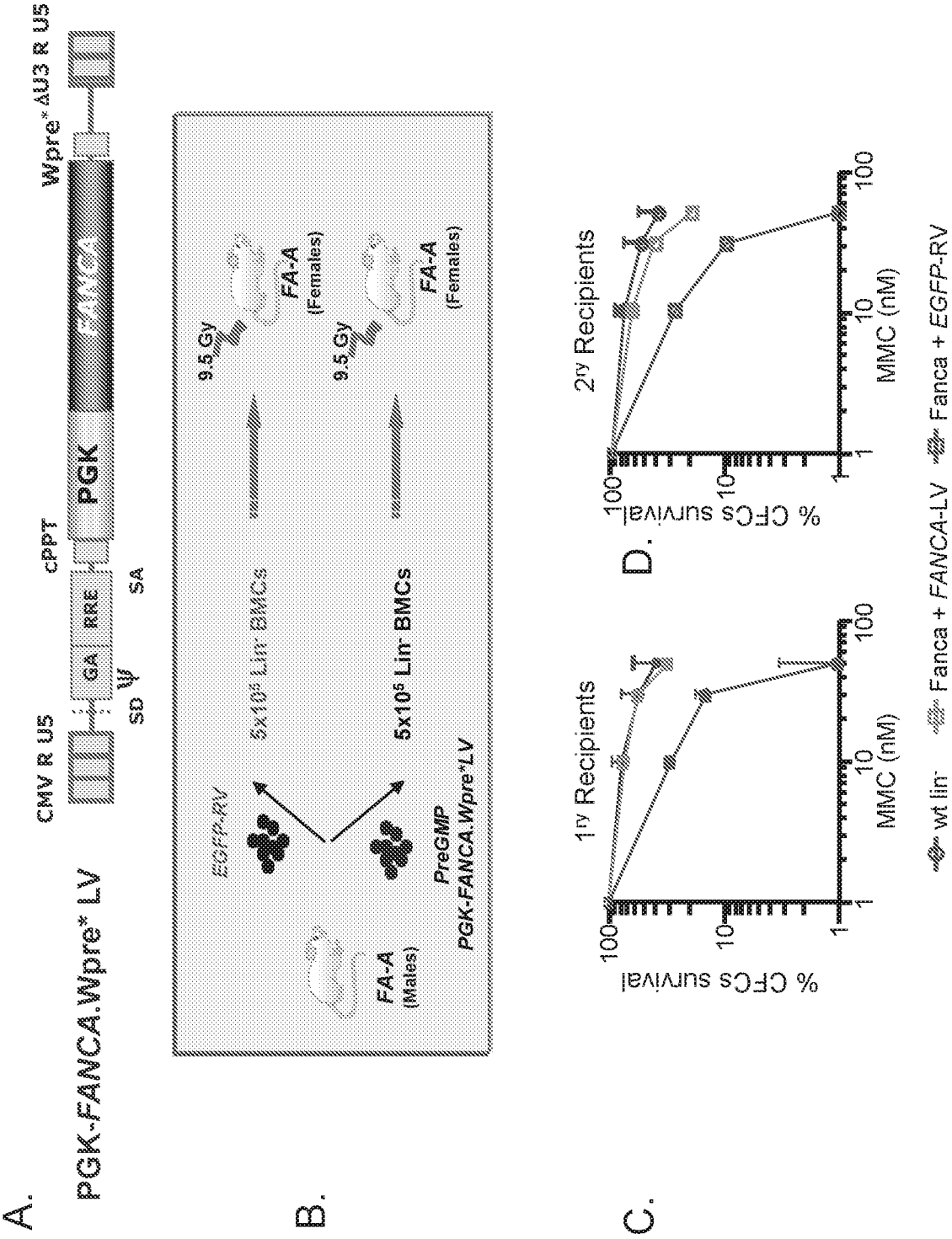


Figure 5.

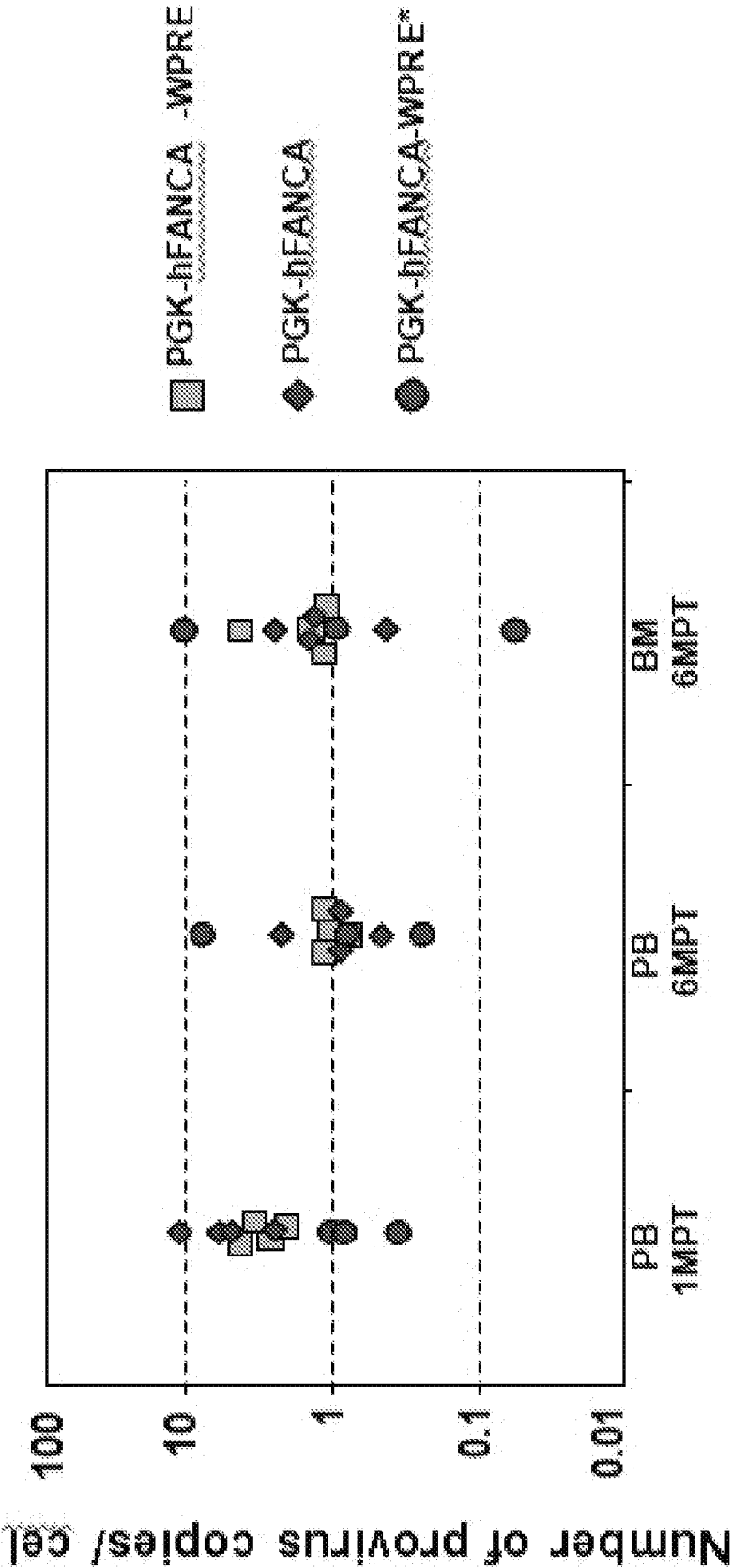


Figure 6.

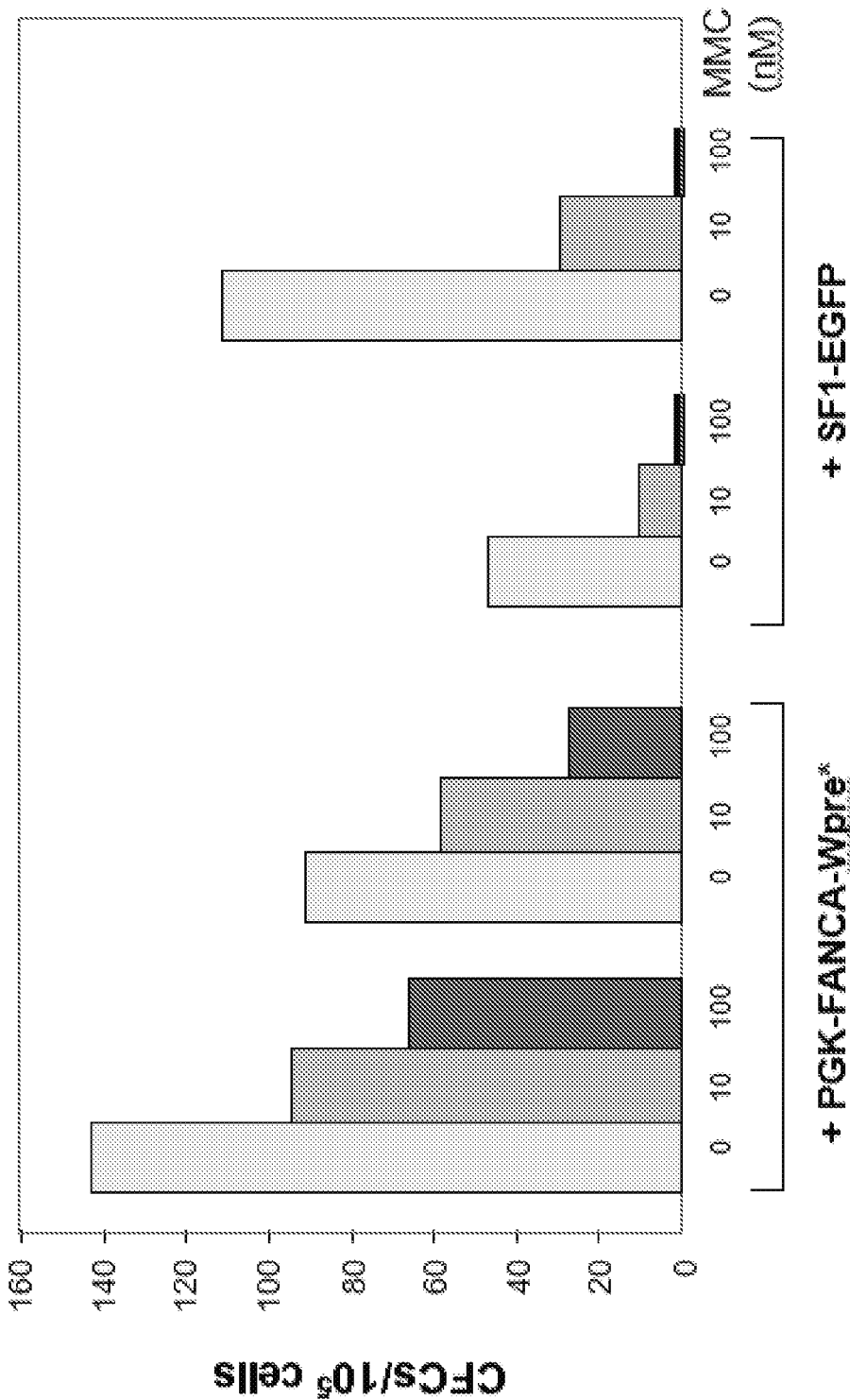


Figure 7.

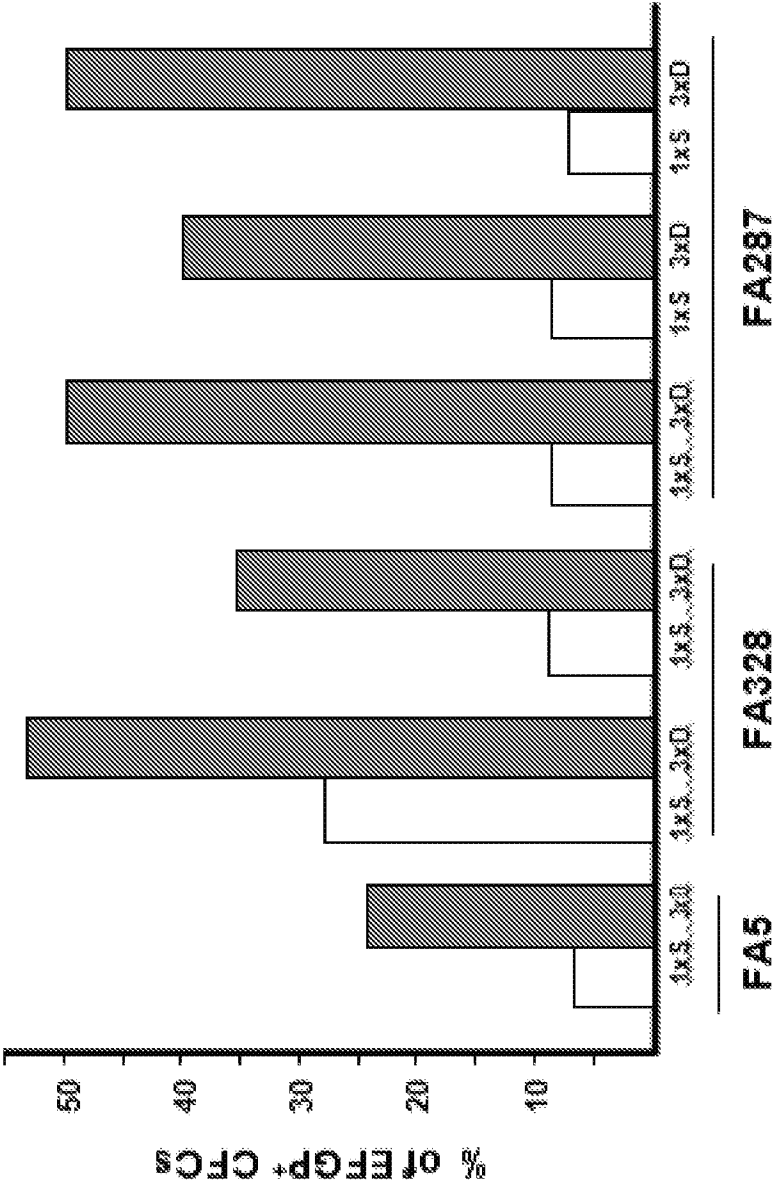


Figure 8.

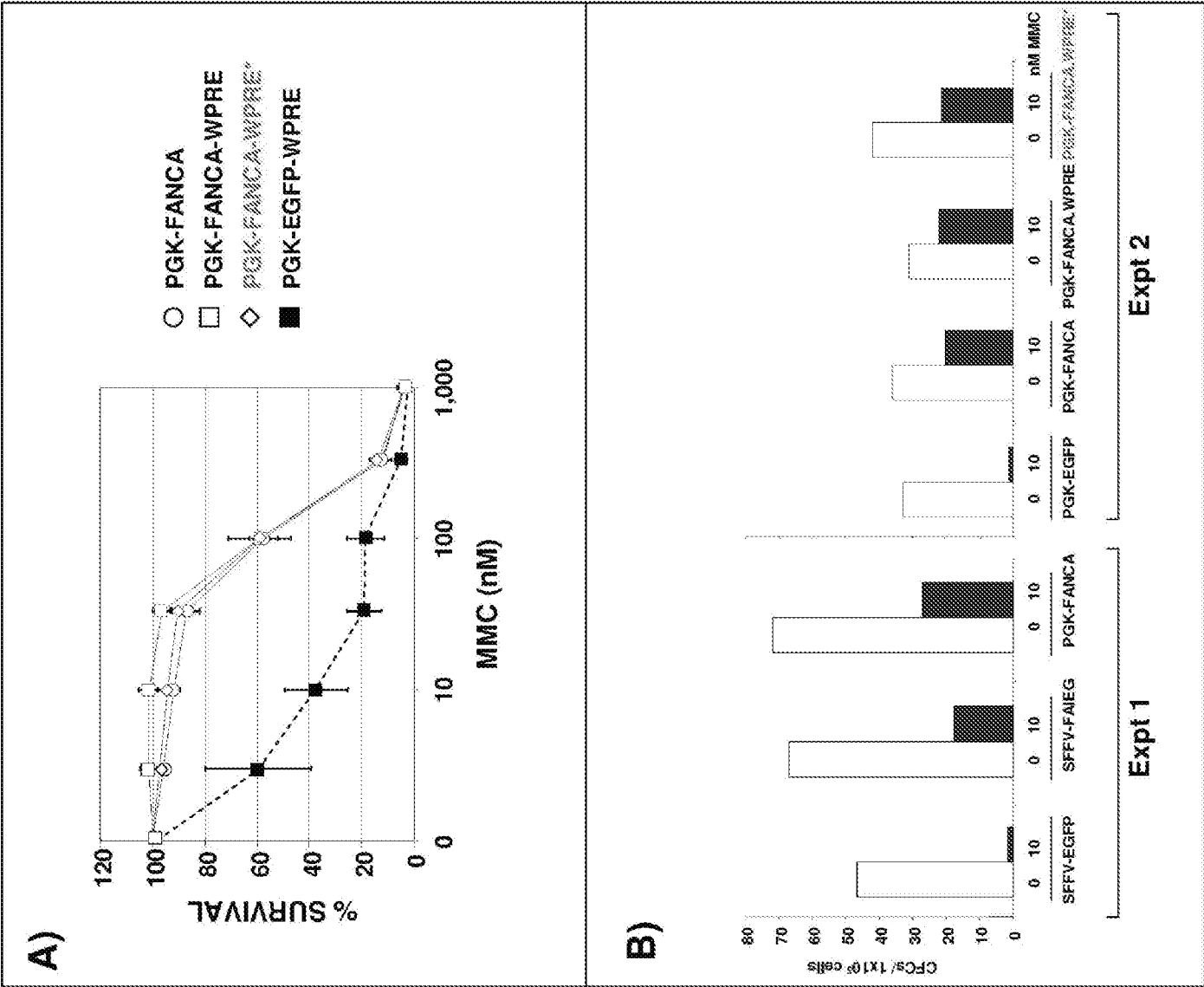


Figure 9.

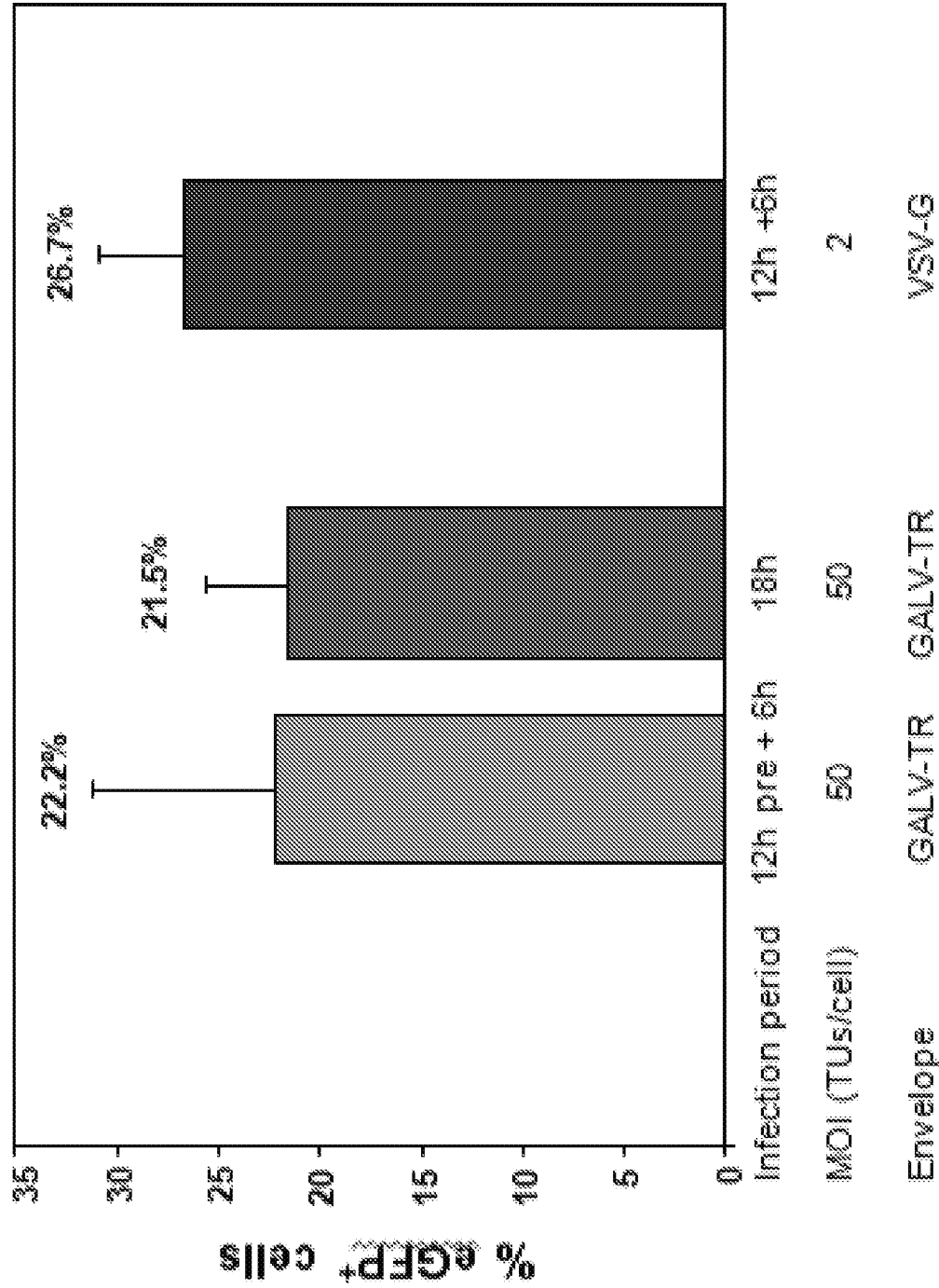


Figure 10.

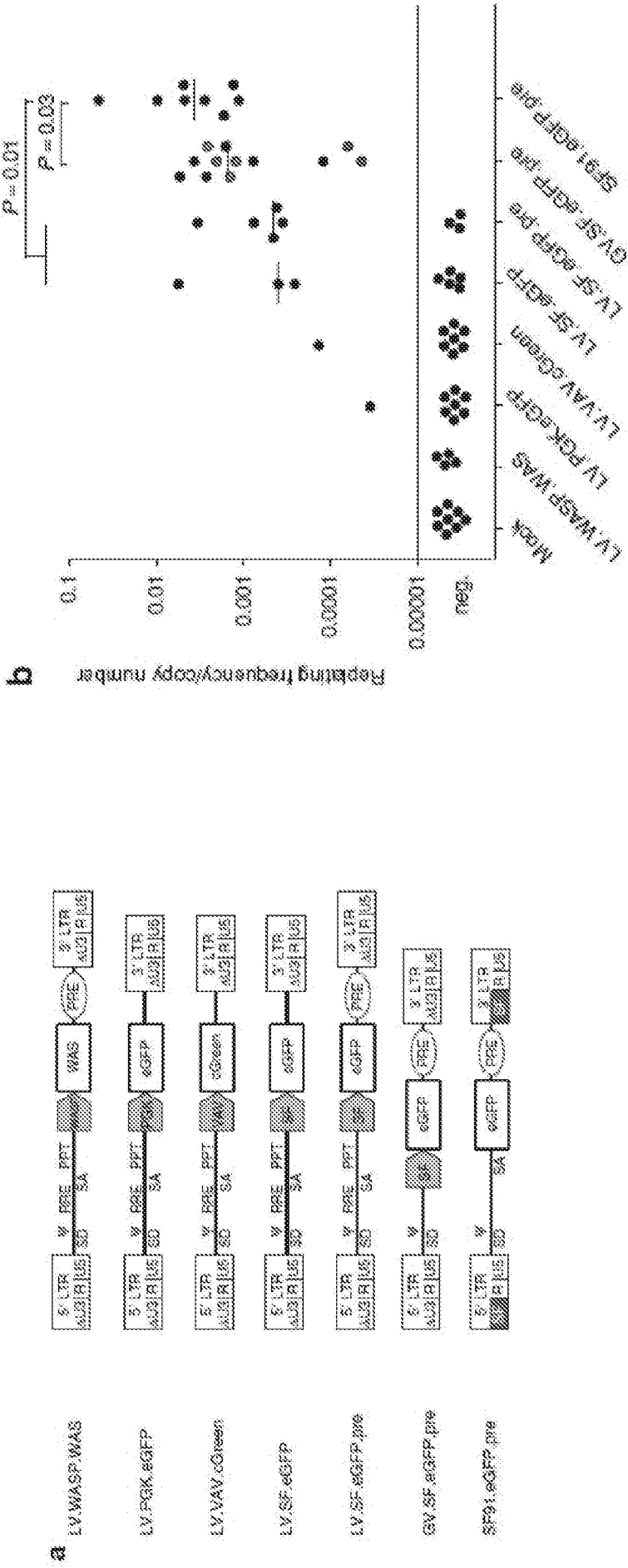


Figure 11.

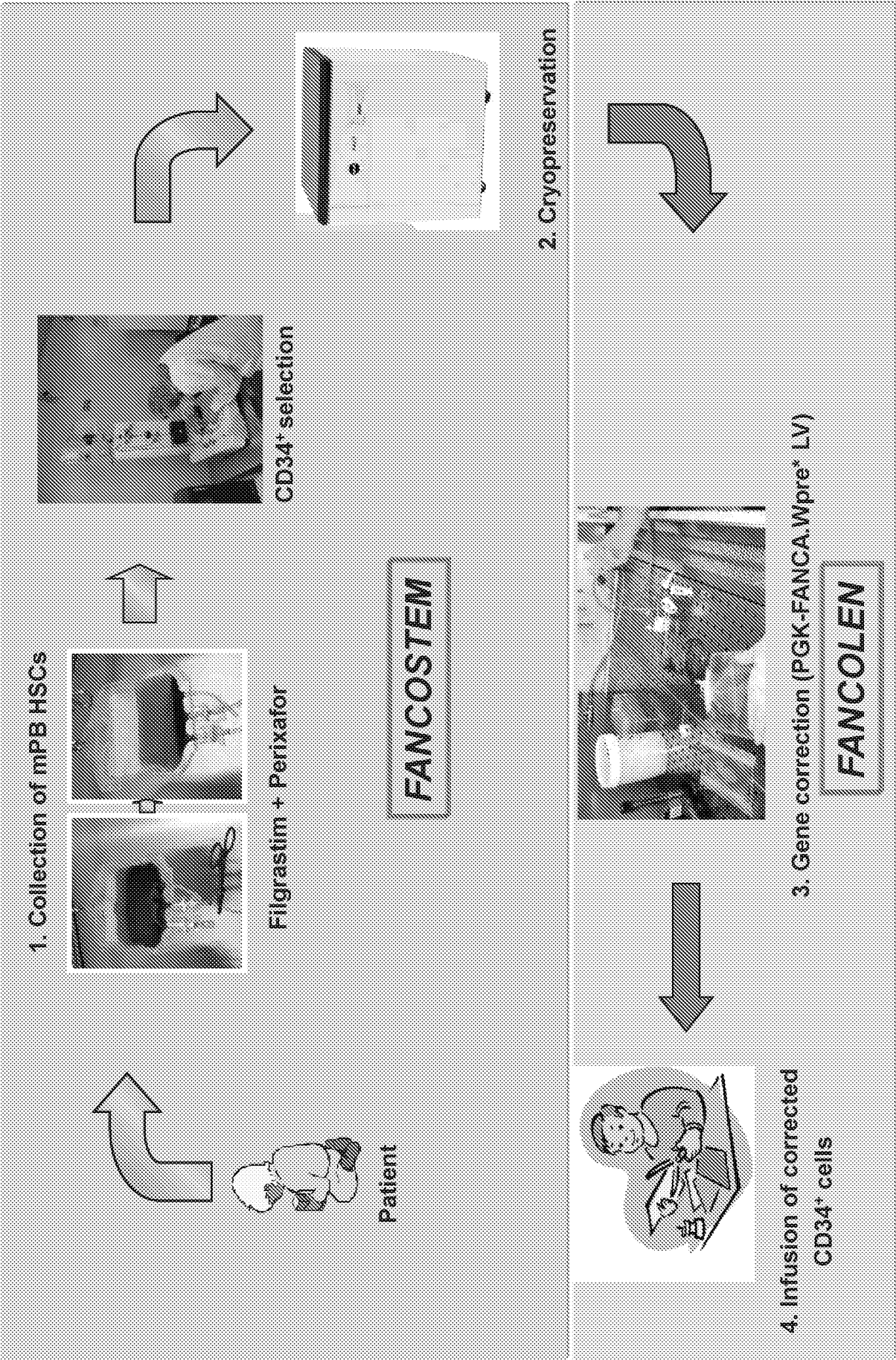


Figure 12.

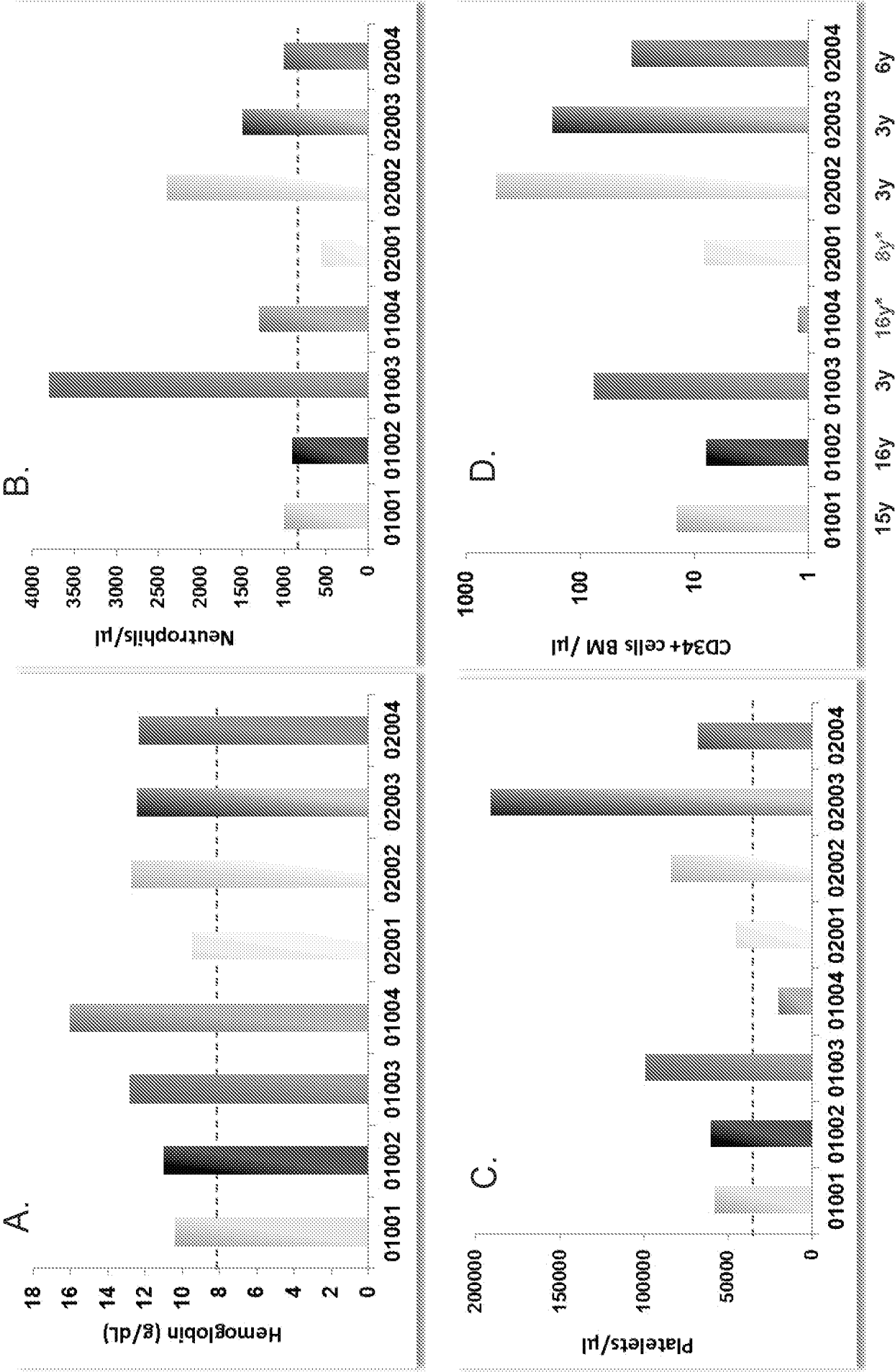


Figure 14.

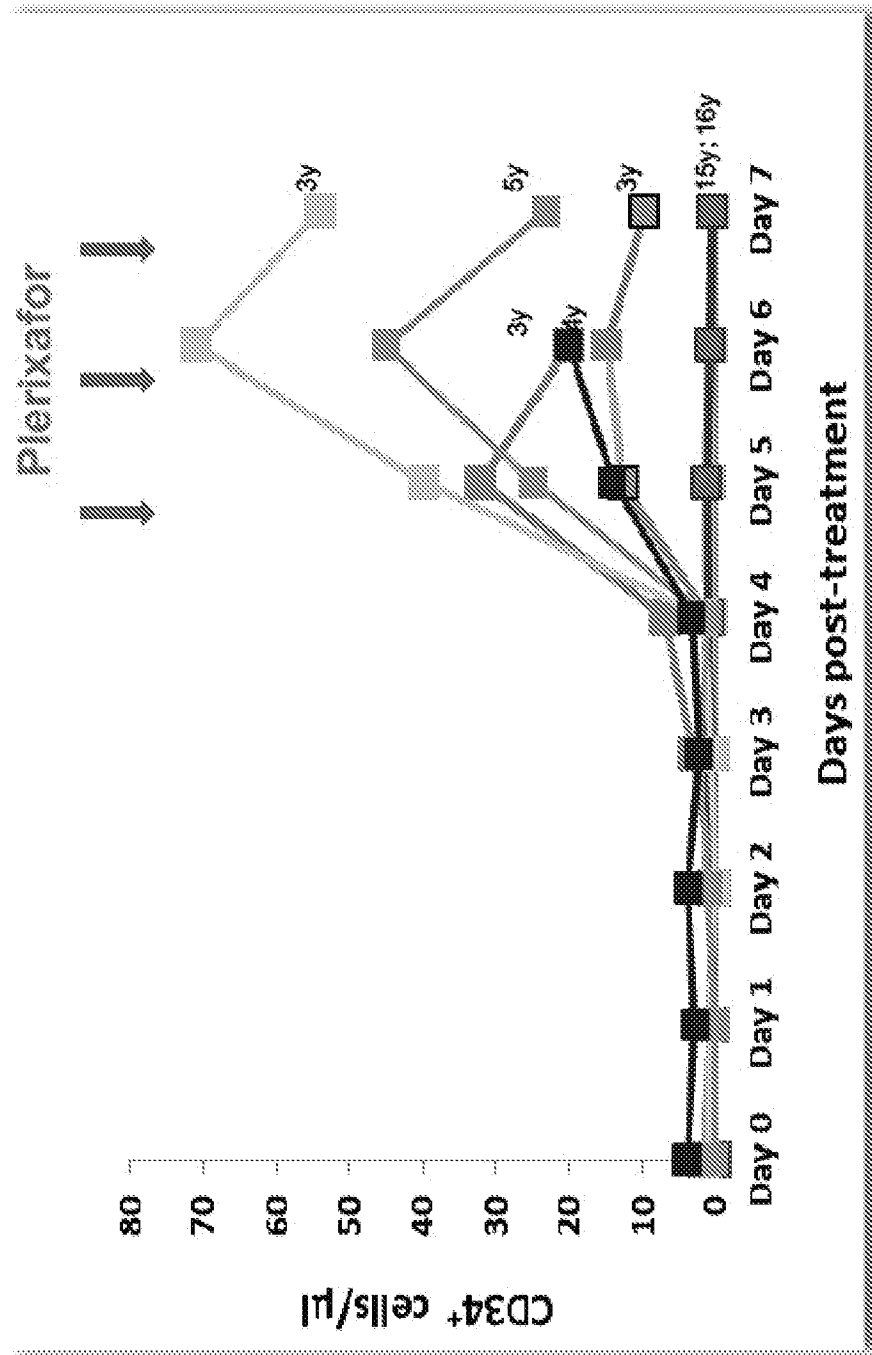


Figure 15.

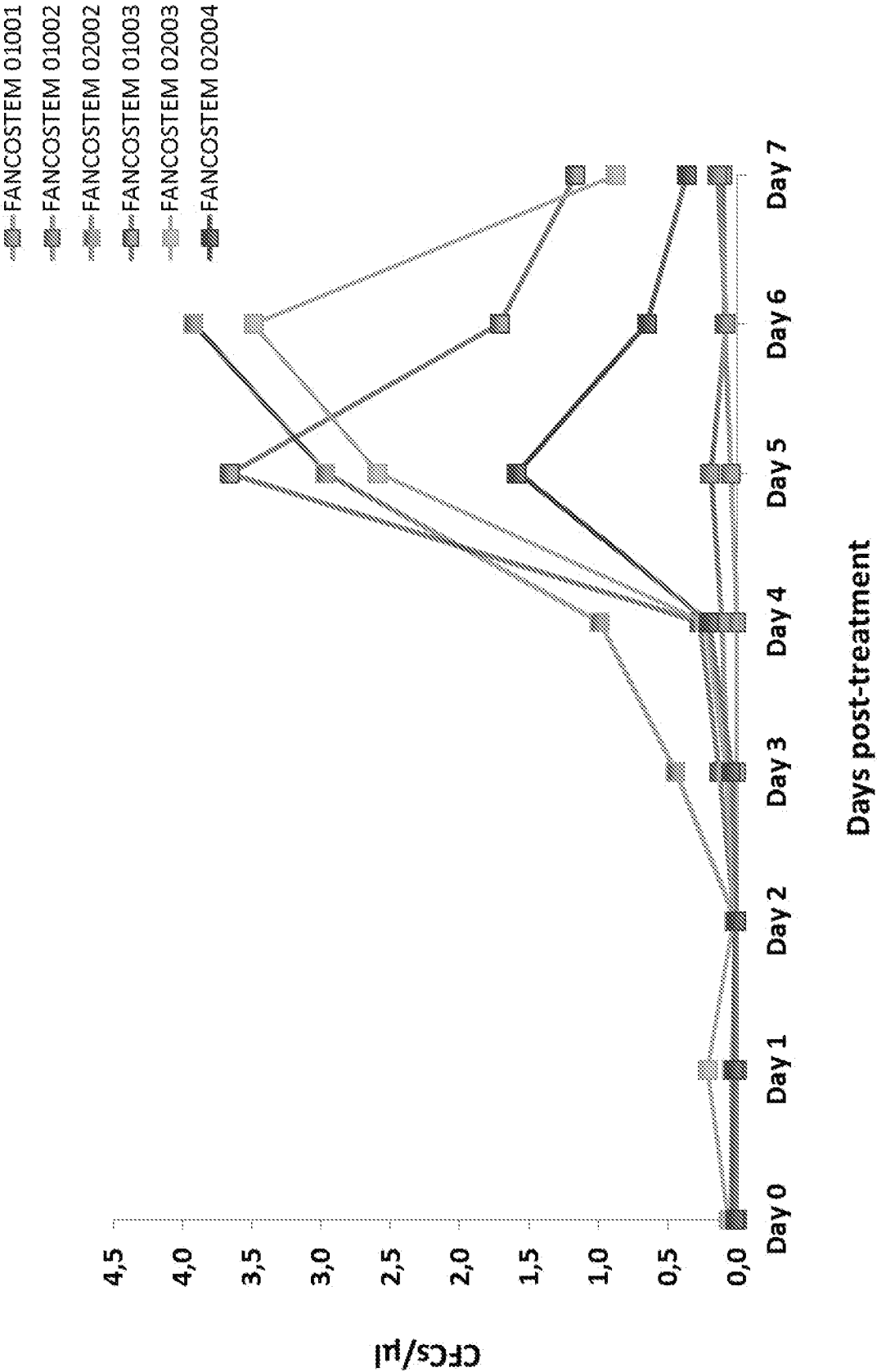
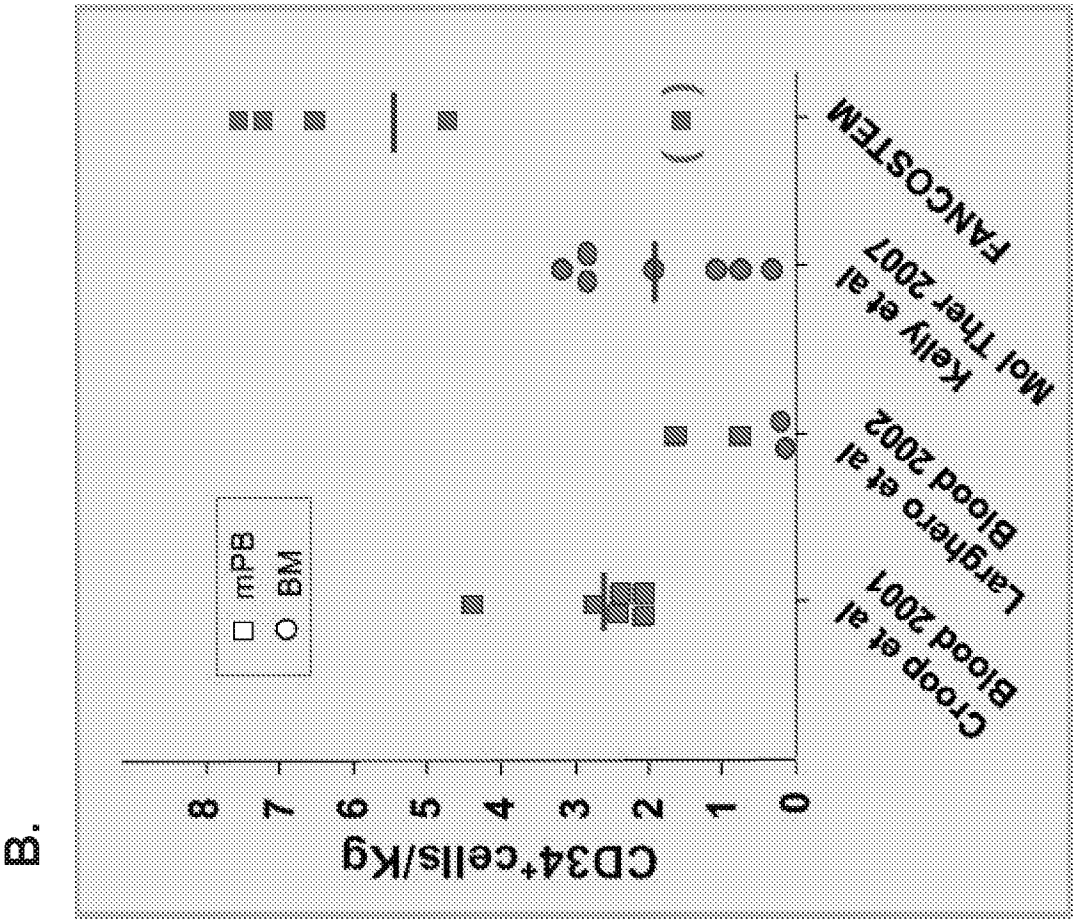


Figure 16.



A.

Figure 17.

Patient	BM			mPB	
	CD34/mcL (x10E6)	CD34/Kg (x10E6)	Corrected CD34/Kg (x10E6)	CD34/Kg (x10E6)	Fold increase vs BM
01003	75	1.5	0.84	1.6	1.9
02002	546	10.9	6.1	7.2	1.2
02003	175	3.5	2.0	7.6	3.9
02004	35	0.7	0.39	6.6	16.6
02005	276	5.5	3.9	4.8	1.5

Figure 18.

	Apheresis		Positive Fraction		
	CD34 ⁺ /Kg	% CD34 ⁺	CD34 ⁺ /Kg	% CD34 ⁺	Recovery
Patient					
1003	1.6x10 ⁶	0.05	8.6x10 ⁵	17.40	53.20
2002	7.2x10 ⁶	0.16	5.1x10 ⁶	79.96	70.94
2003	7.6x10 ⁶	0.17	2.9x10 ⁶	29.33	47.85
2004	6.6x10 ⁶	0.15	2.0x10 ⁶	25.55	26.63

Figure 19.

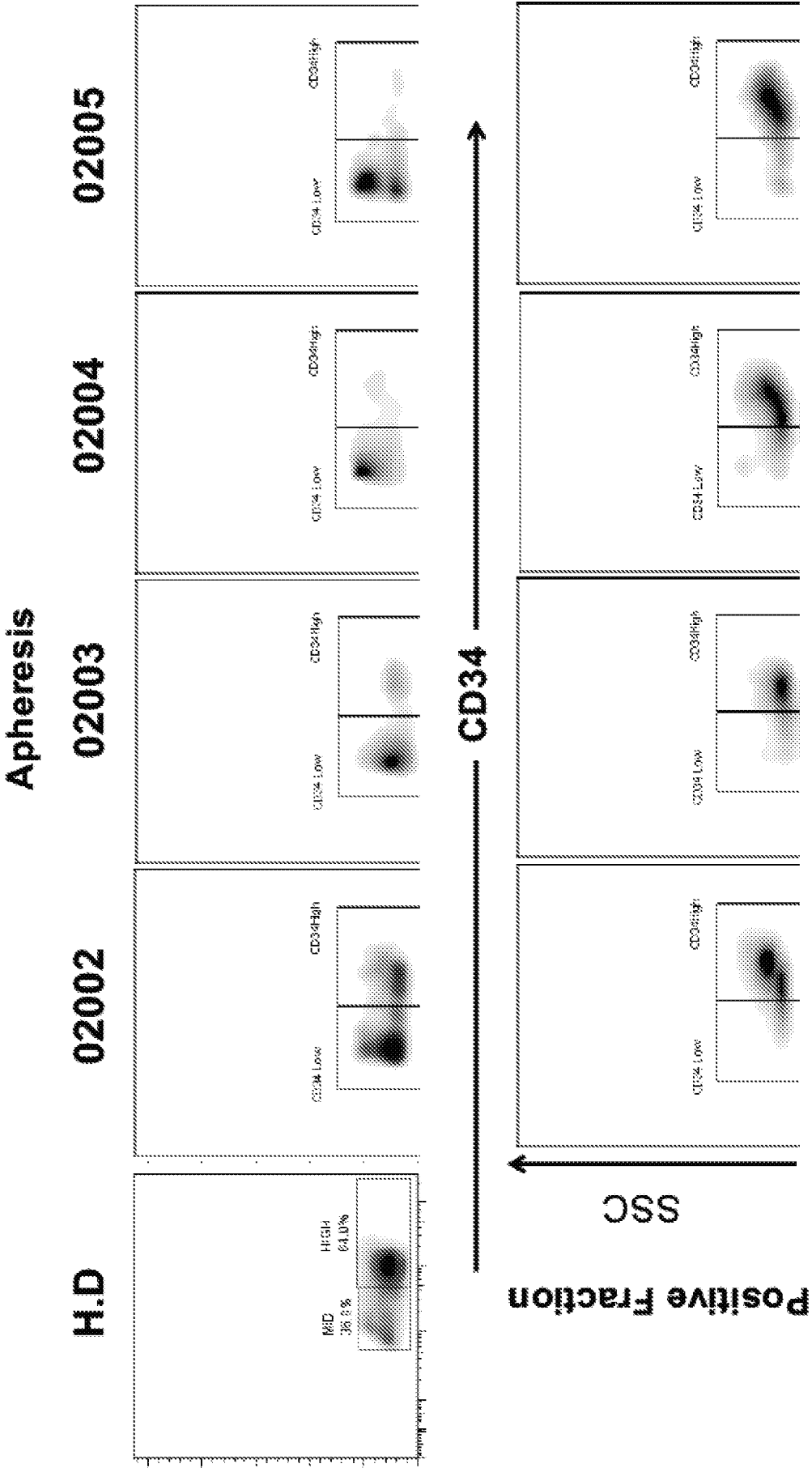
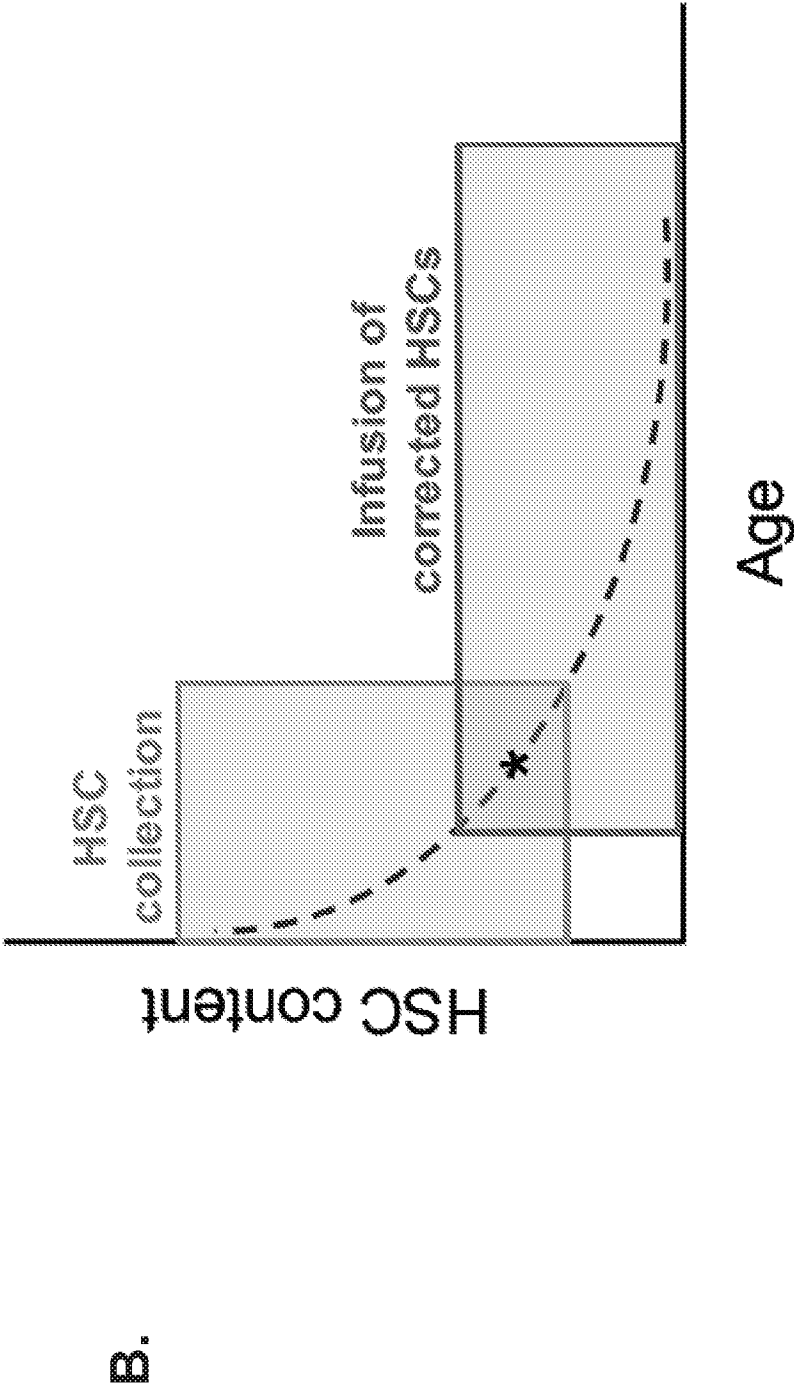


Figure 20.

A.

PB cell counts: Hb (g/dL)		Neutroph. (N°/ μ L)		Platelets: (N°/ μ L)	
FA 02005:	12.5	1,680		38,000	
FANCOSTEM:	>8.0	>1,000	or	>30,000	
FANCOLEN:	<8.0	<1,000	or	<50,000	



Not mosaic; Homozygote FANCA c.239 C>T p.Gln99*

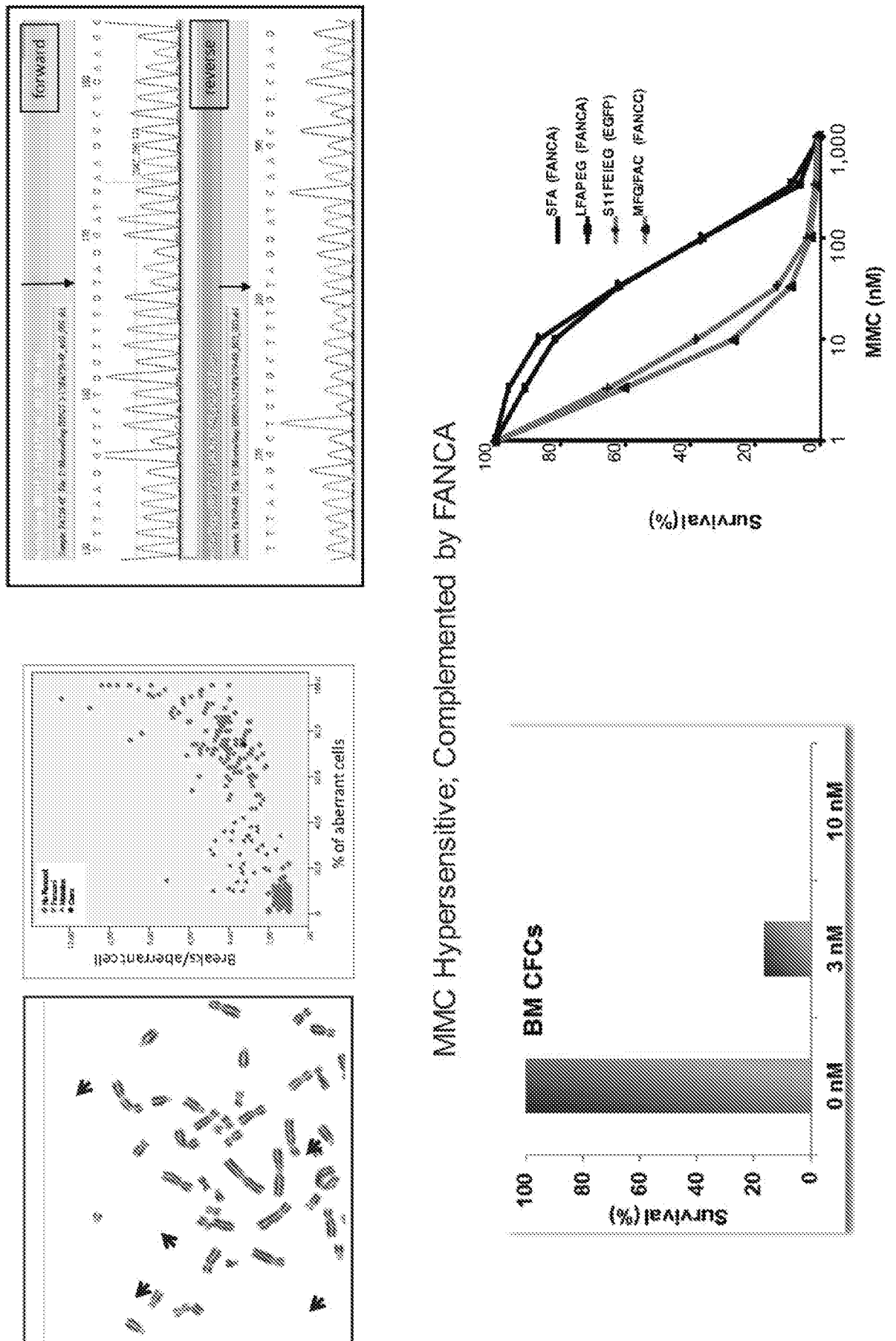


Figure 22.

02005	Total CD34 ⁺ cells	CD34 ⁺ /Kg	% CD34 ⁺	Viability	Total CFCs	CFCs/Kg
Pre- transduction	1.86x10 ⁷	1.02 x10 ⁶	59.0%	100 %	2.9 x10 ⁵	1.7x10 ⁴
Post- transduction	2.3x10 ⁷	1.33 x10 ⁶	80.5%	97.4 %	2.7 x10 ⁵	1.5x10 ⁴
Yield	125.2%				110.6%	

Figure 23.

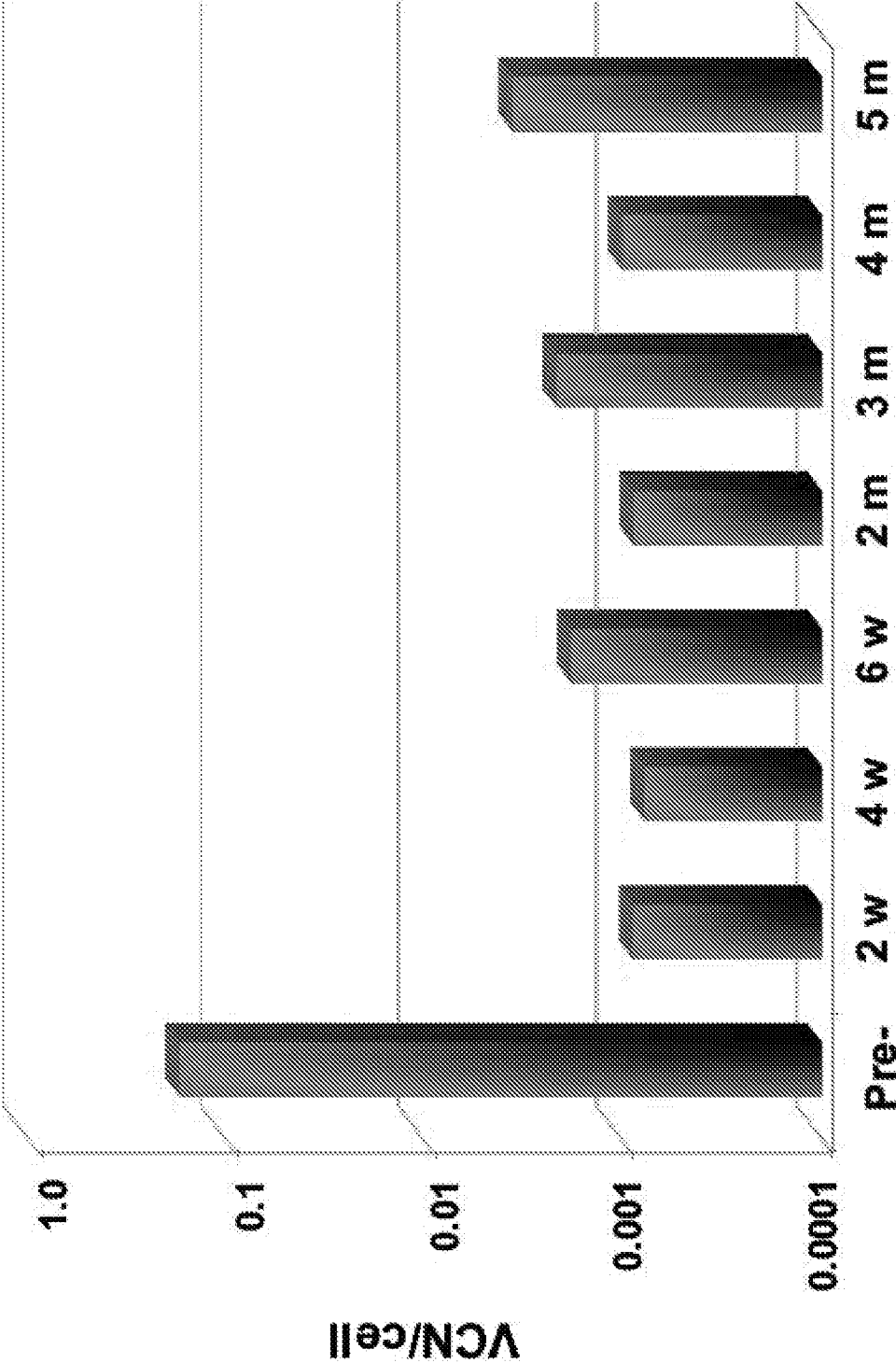


Figure 24.

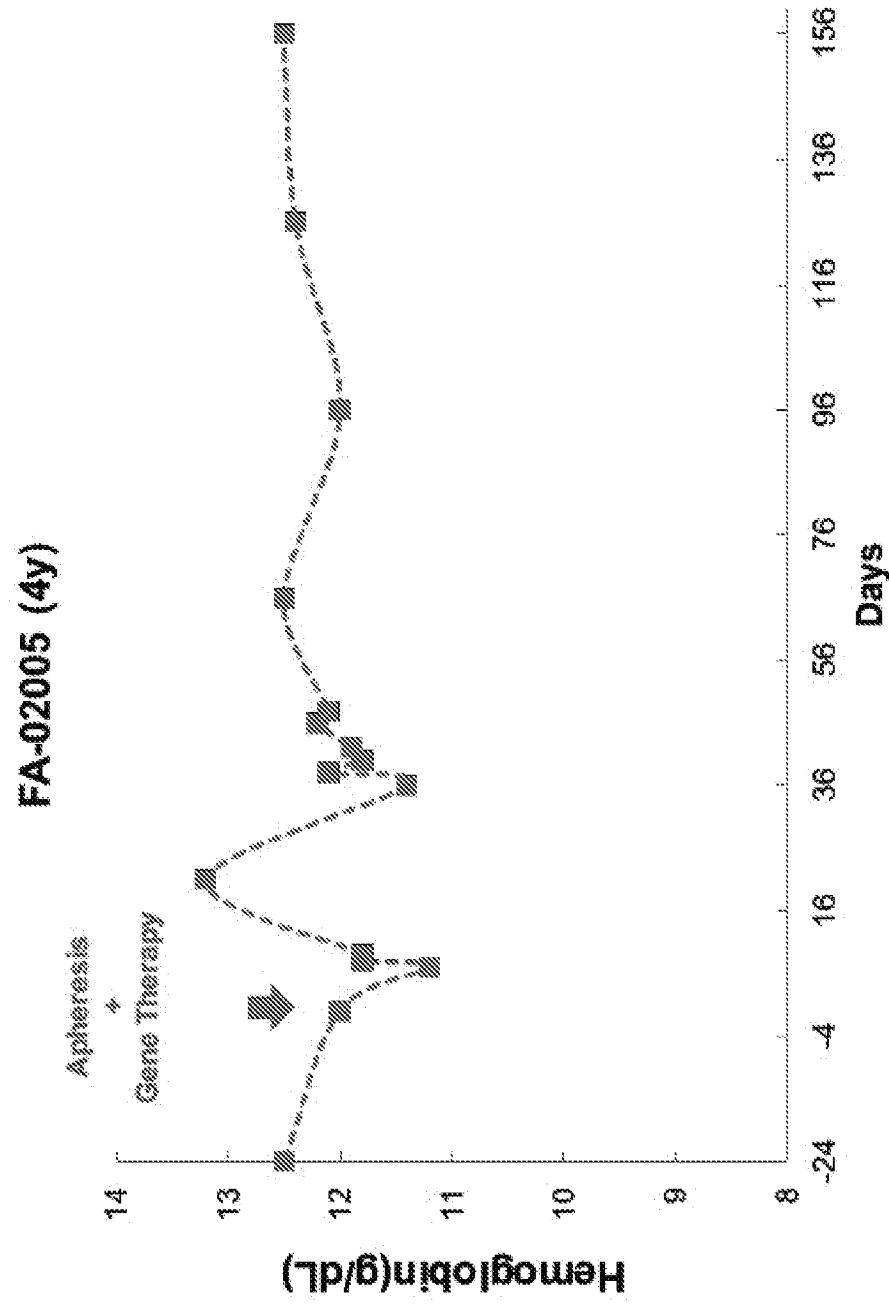


Figure 25.

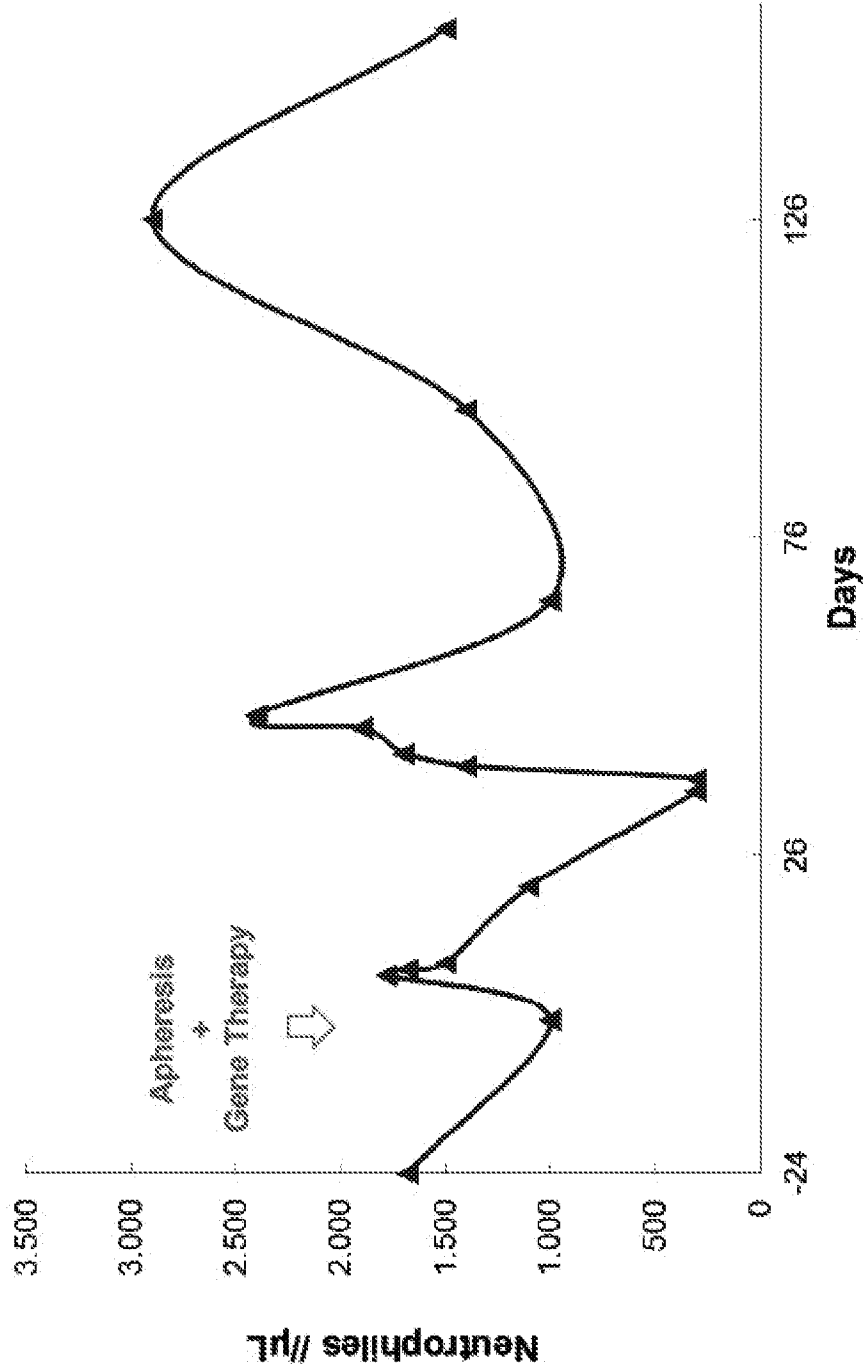


Figure 26.

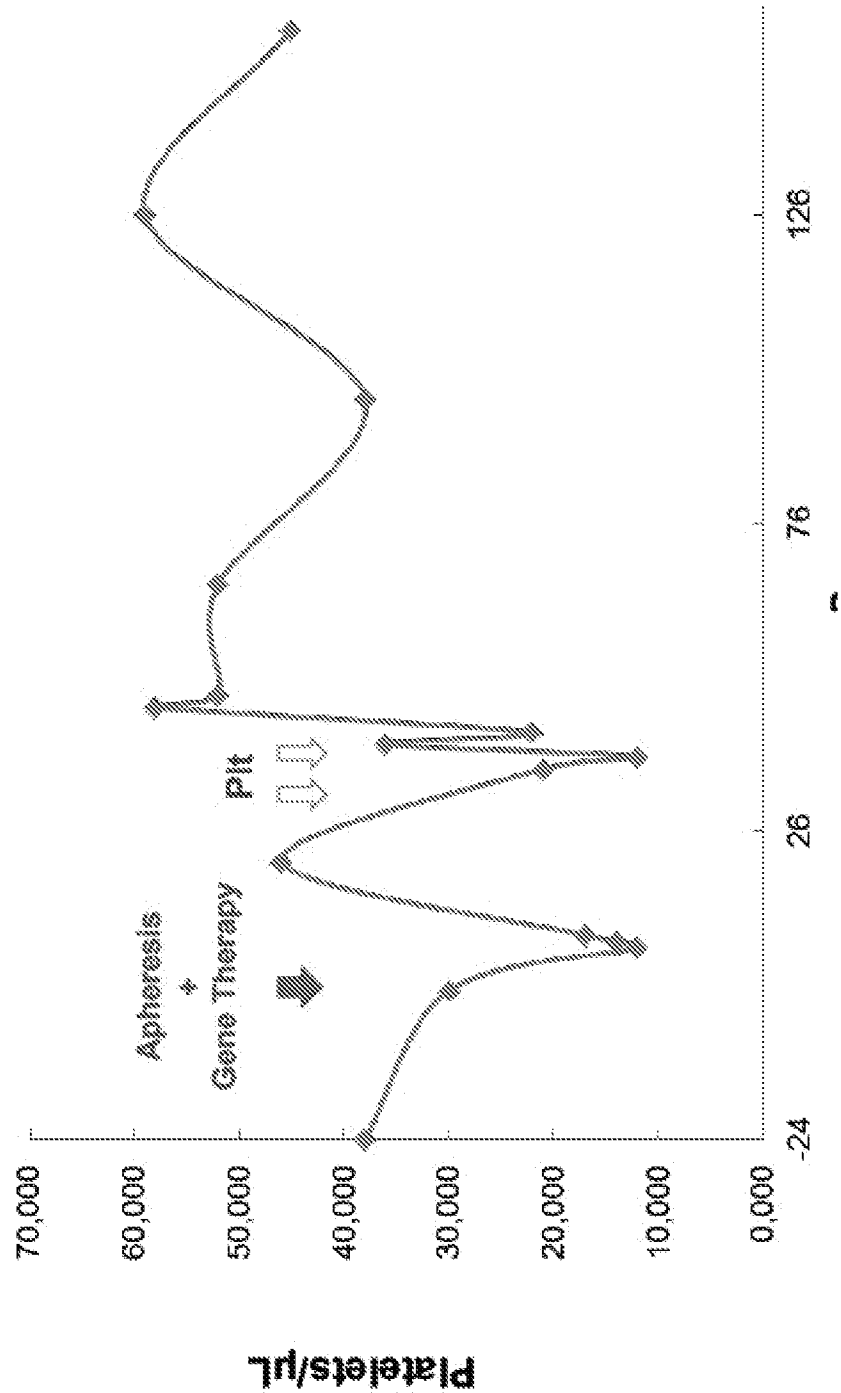


Figure 27.

12/5/2014				
PB cell counts: Hb (g/dL)		Neutroph. (N°/ µL)		Platelets: (N°/ µL)
FA 02002:	12.7	2,400		84,000
FANCOSTEM:	>8.0	or	>1,000	or >30,000
FANCOLEN:	<8.0	or	<1,000	or <50,000
20/1/2016				
PB cell counts: Hb (g/dL)		Neutroph. (N°/ µL)		Platelets: (N°/ µL)
FA 02002:	10.5	1,600		29,000
FANCOLEN:	<8.0	or	<1,000	or <50,000

Figure 28.

Not mosaic; Homozygote *FANCA* c.239 C>T p.Gln99*
MMC Hypersensitive; Complemented by *FANCA*

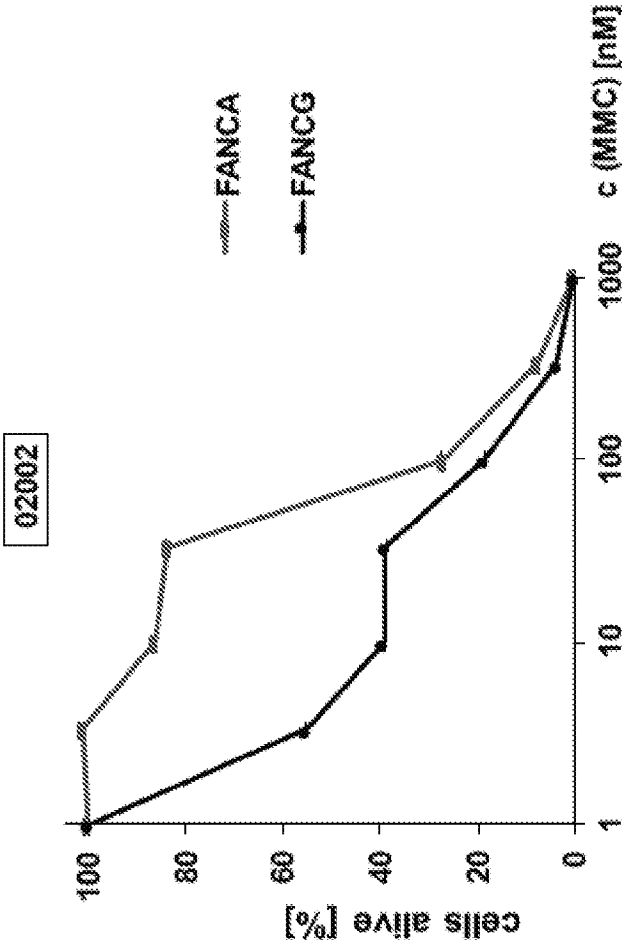


Figure 29.

02002	Total CD34 ⁺ cells	CD34 ⁺ /Kg	% CD34 ⁺	Viability	Total CFCs	CFCs/Kg
Pre- transduction	1.81x10 ⁷	8.5x10 ⁵	87.6%	92.7%	8.8 x10 ⁵	4.1x10 ⁴
Post- transduction	1.17x10 ⁷	5.5x10 ⁵	90.6%	84.04 %	6.75 x10 ⁵	3.8x10 ⁴
Yield	64.24%				76.7%	

Figure 30.

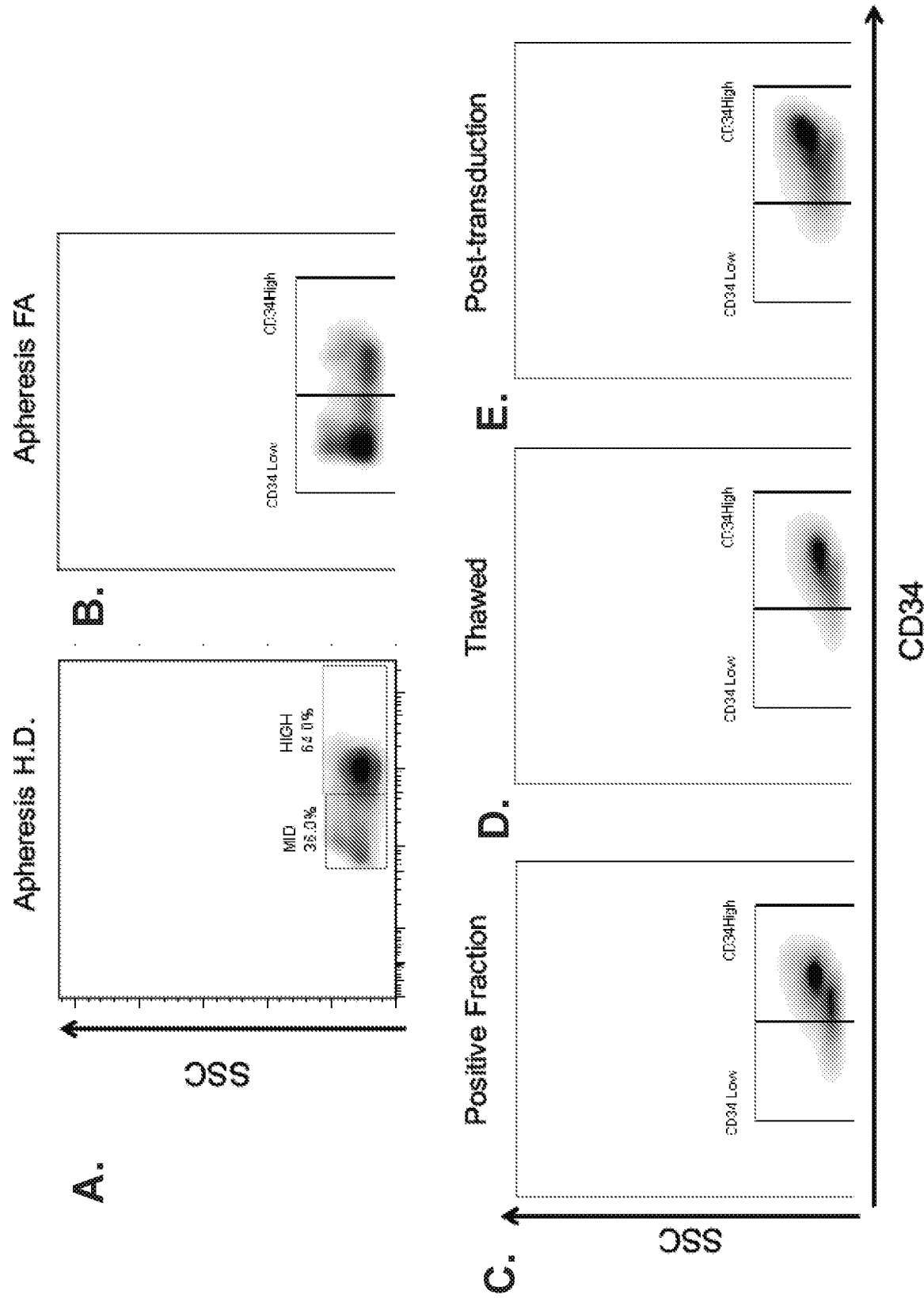


Figure 31.

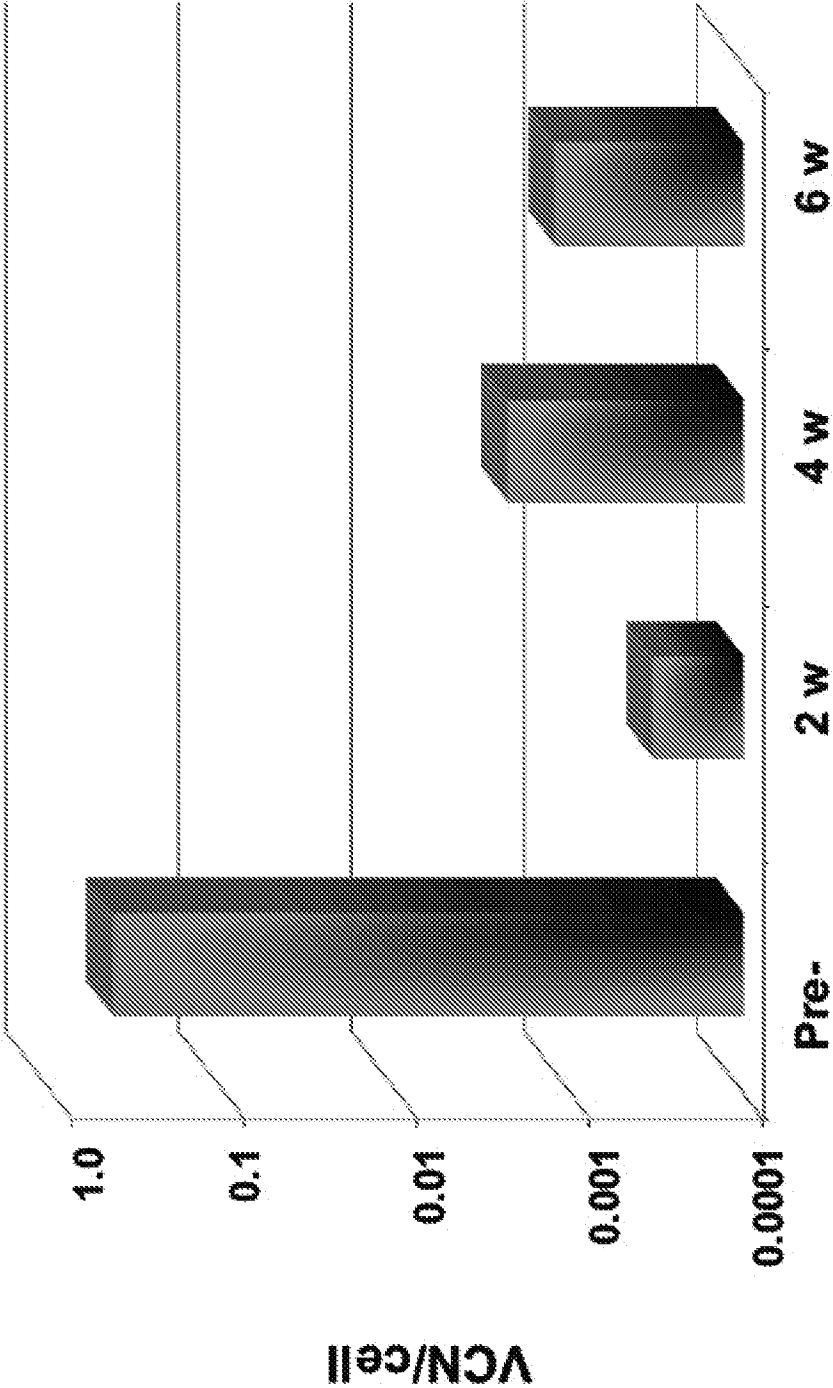


Figure 32.

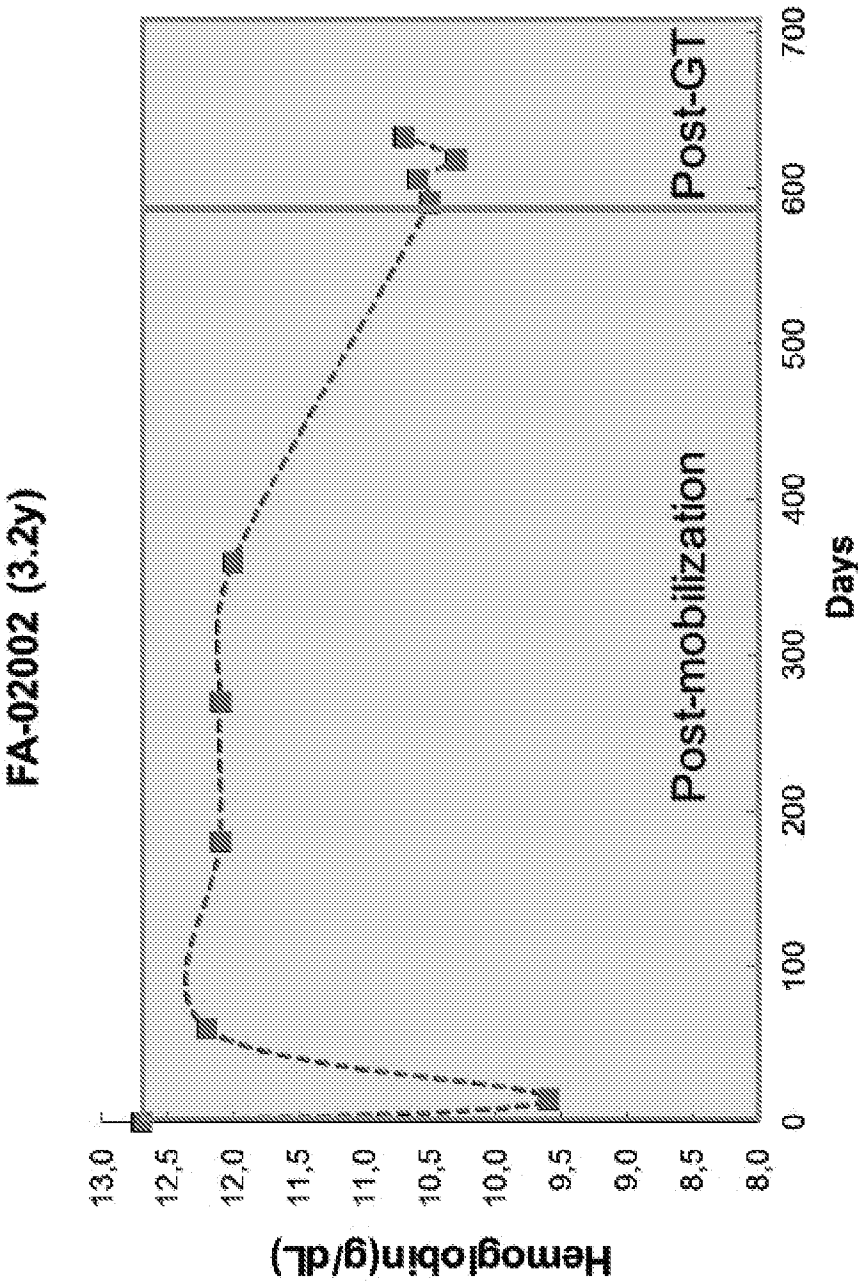


Figure 33.

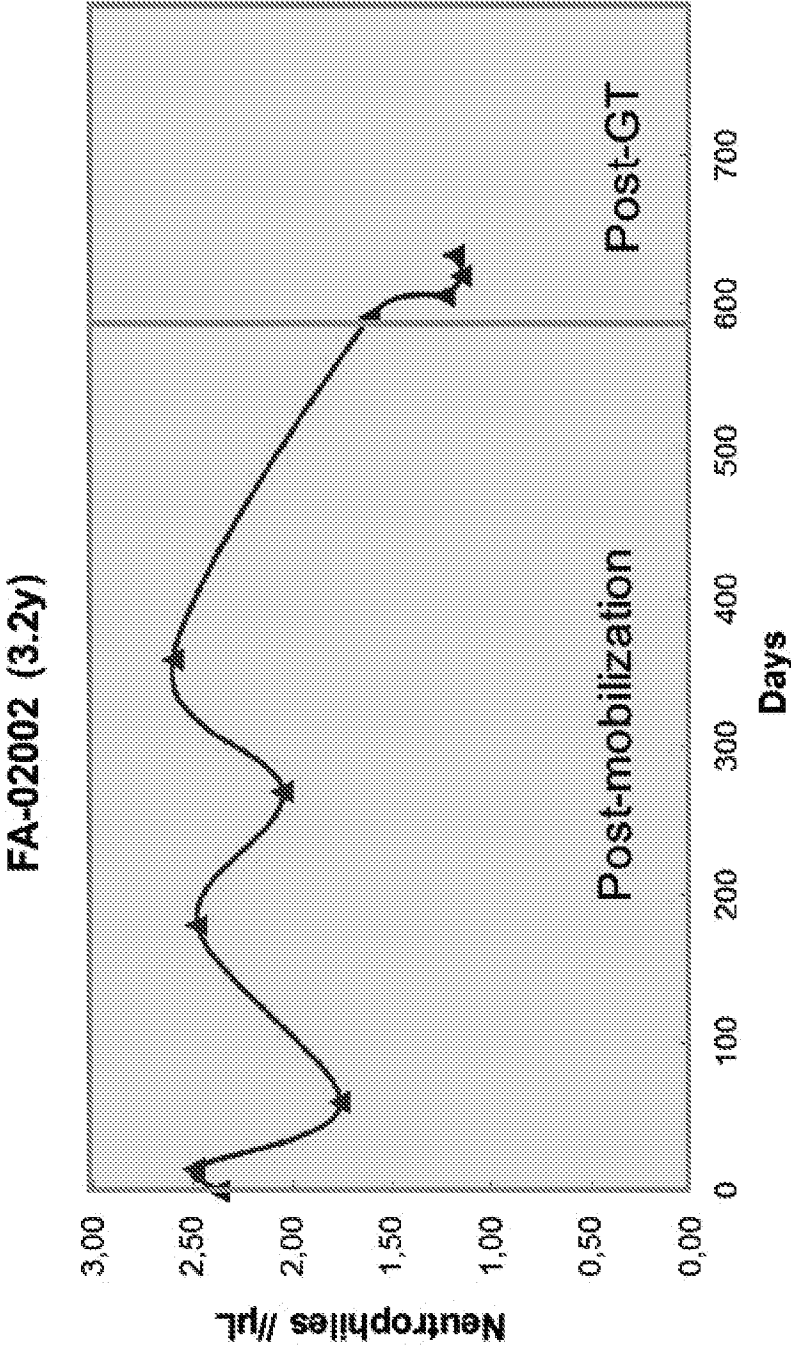


Figure 34.

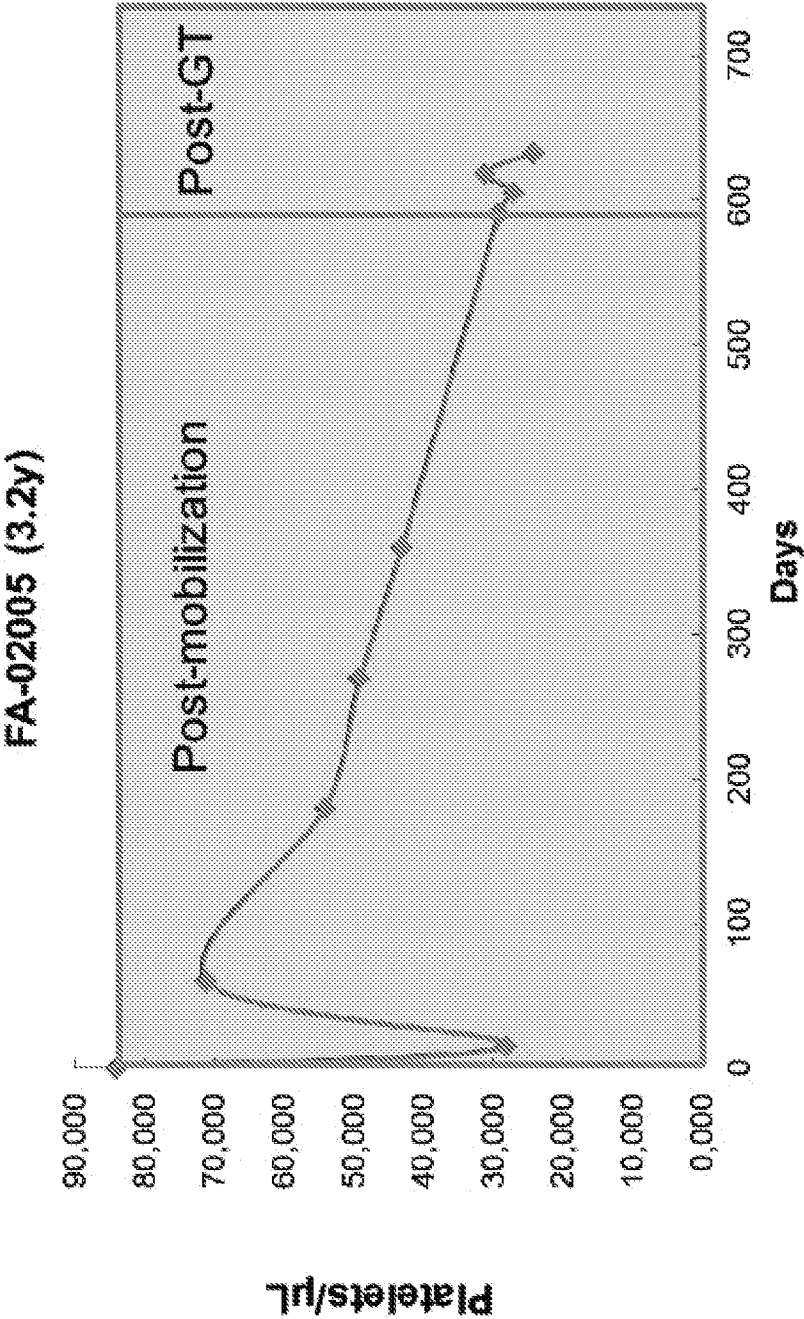


Figure 35.

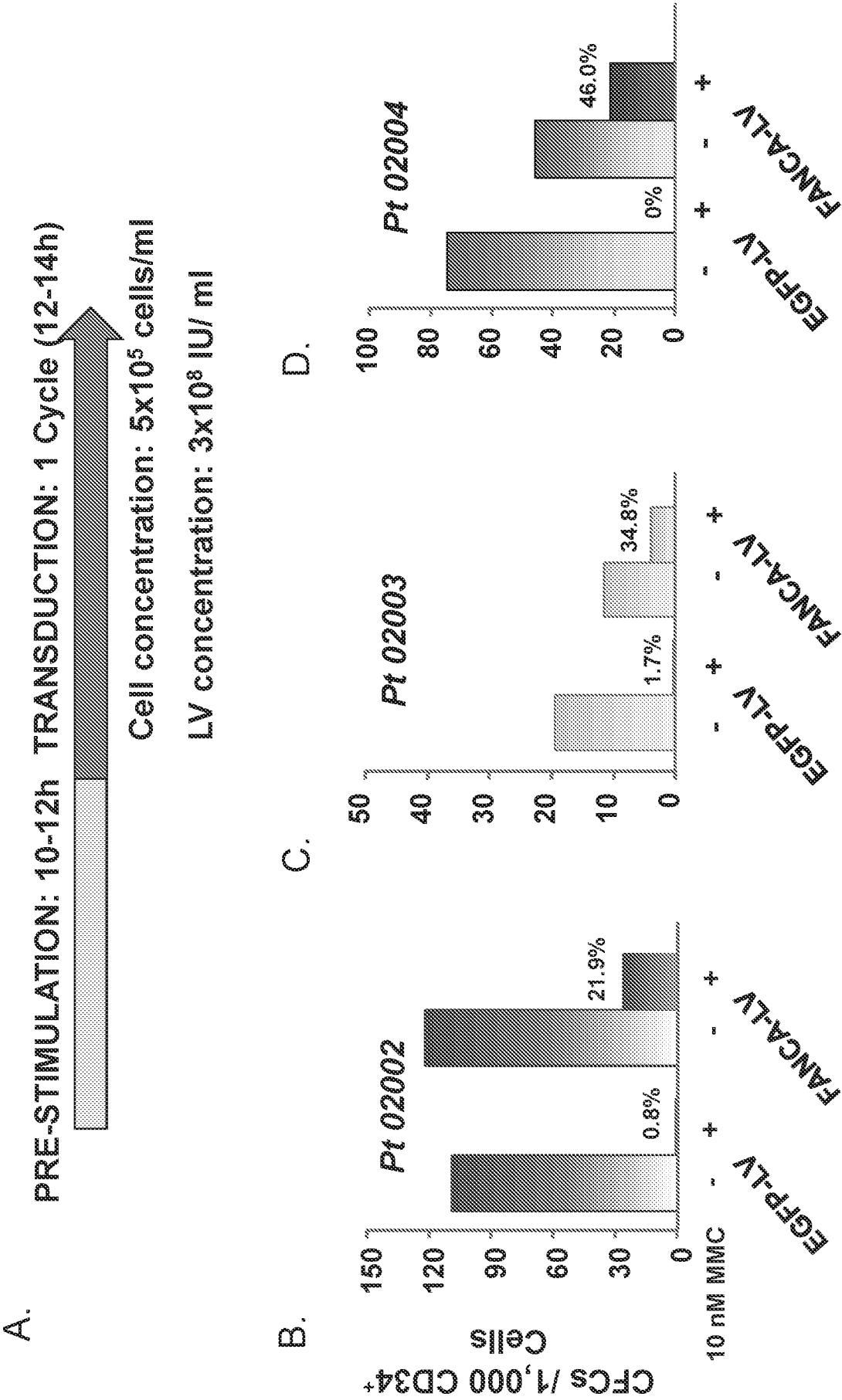
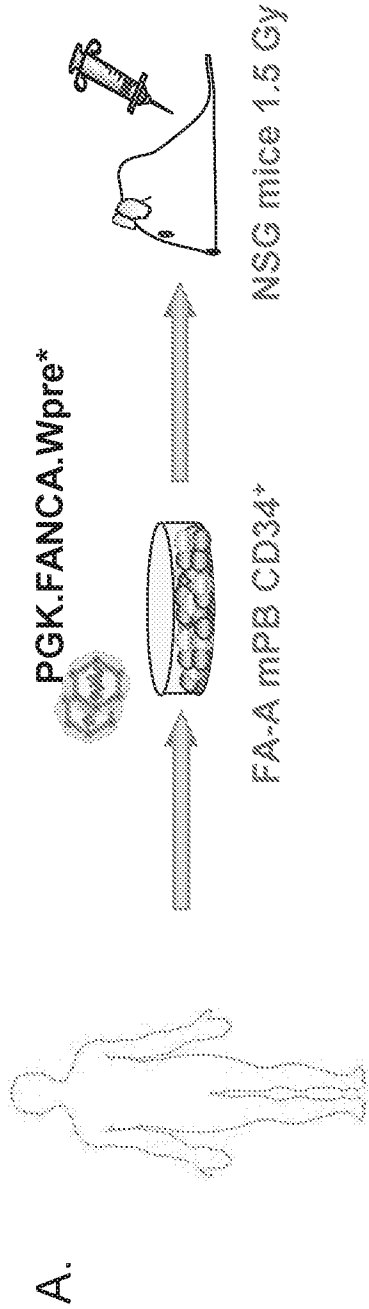


Figure 36.



30 Days post-transplantation

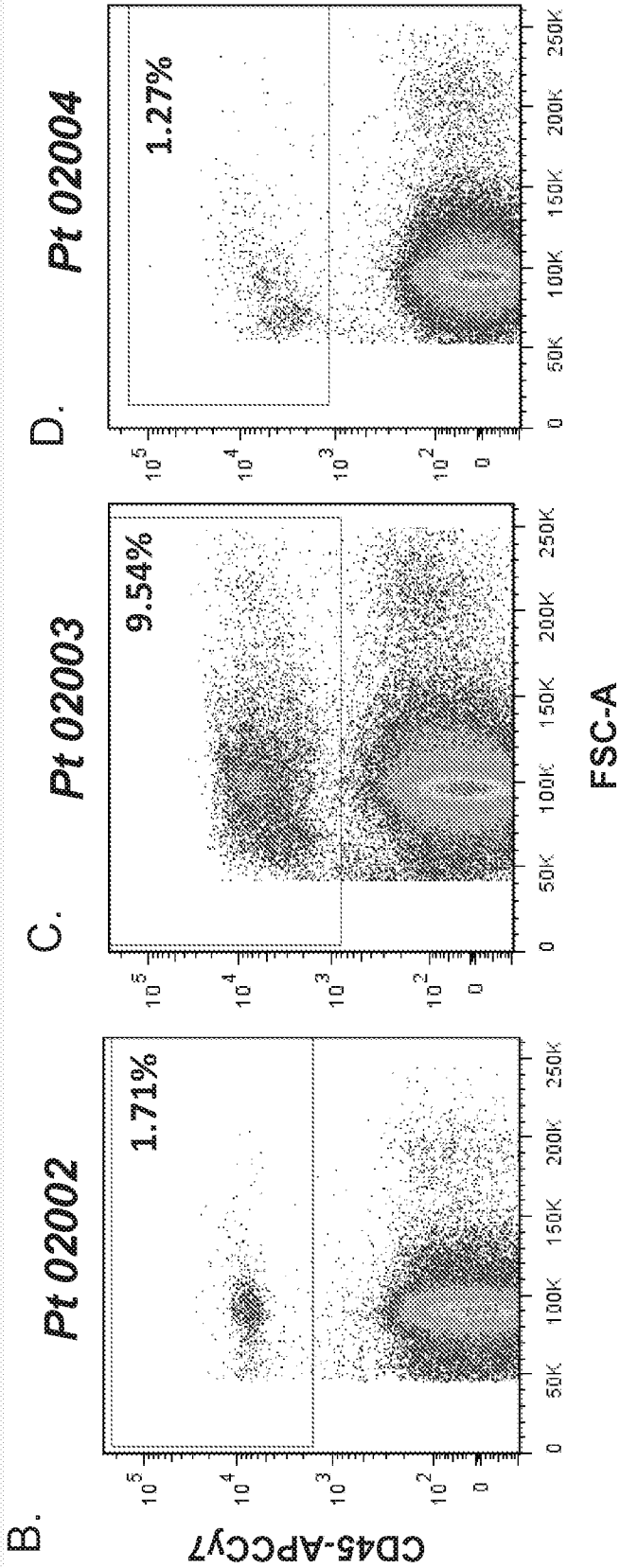


Figure 37.

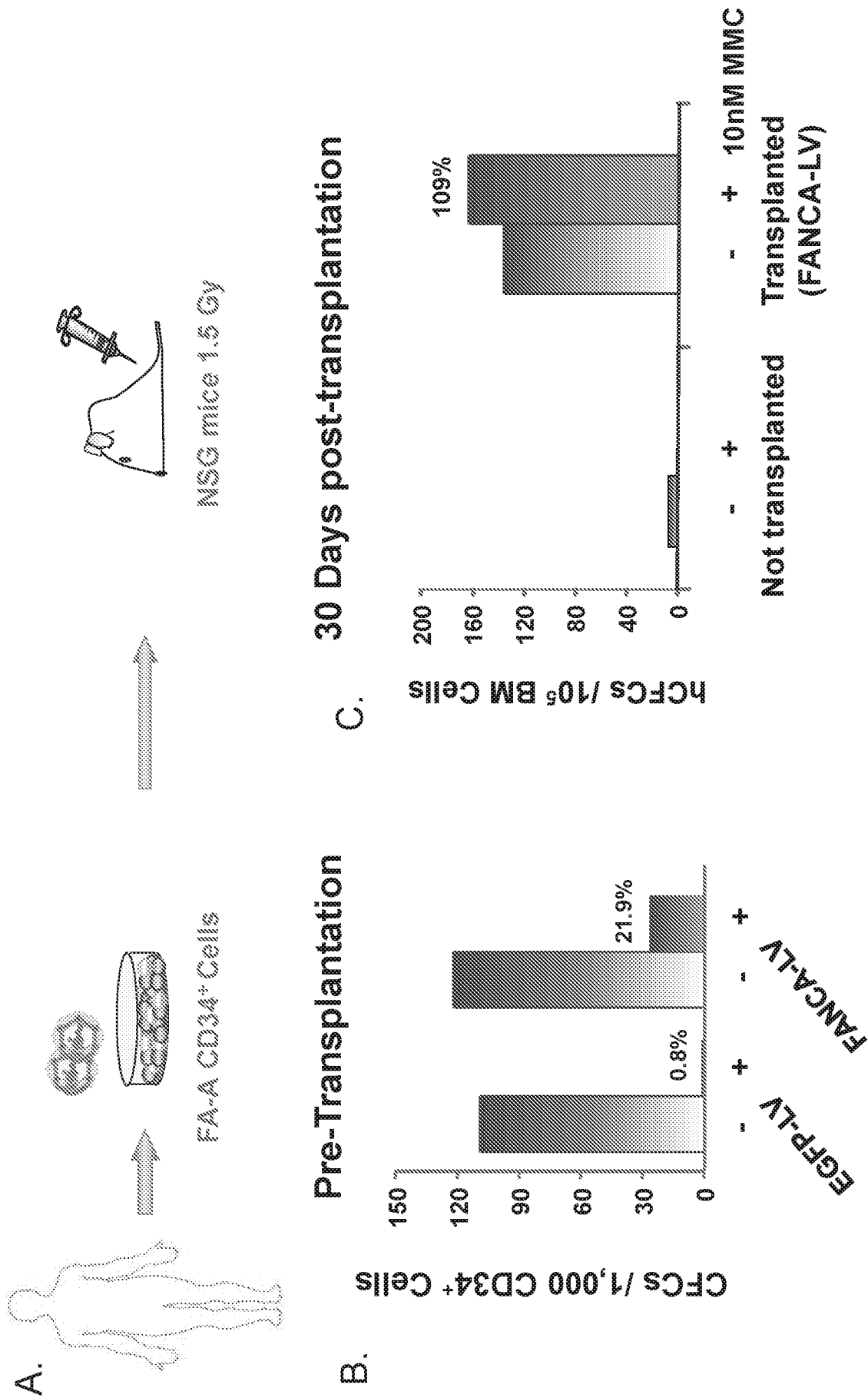


Figure 38.

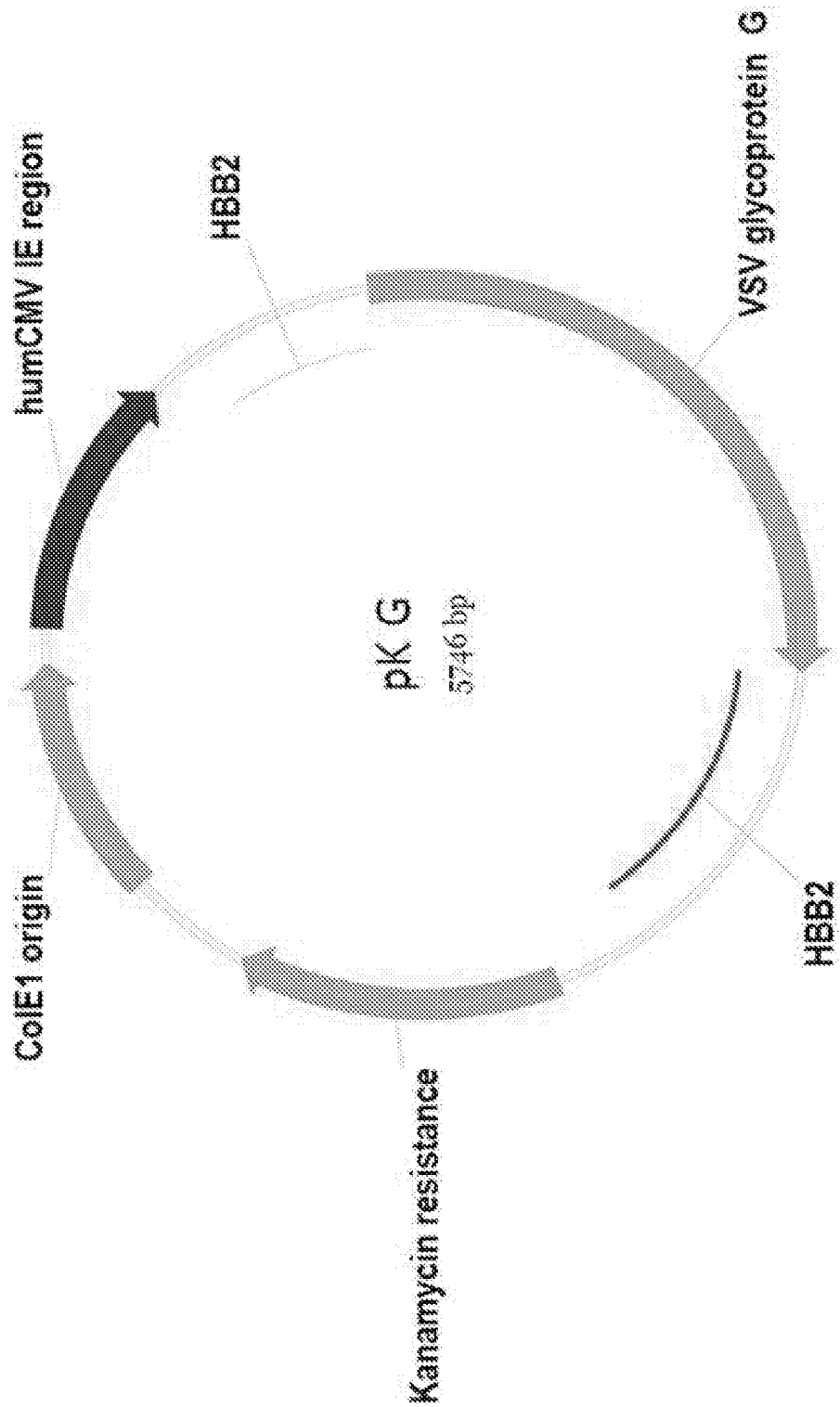


Figure 39.

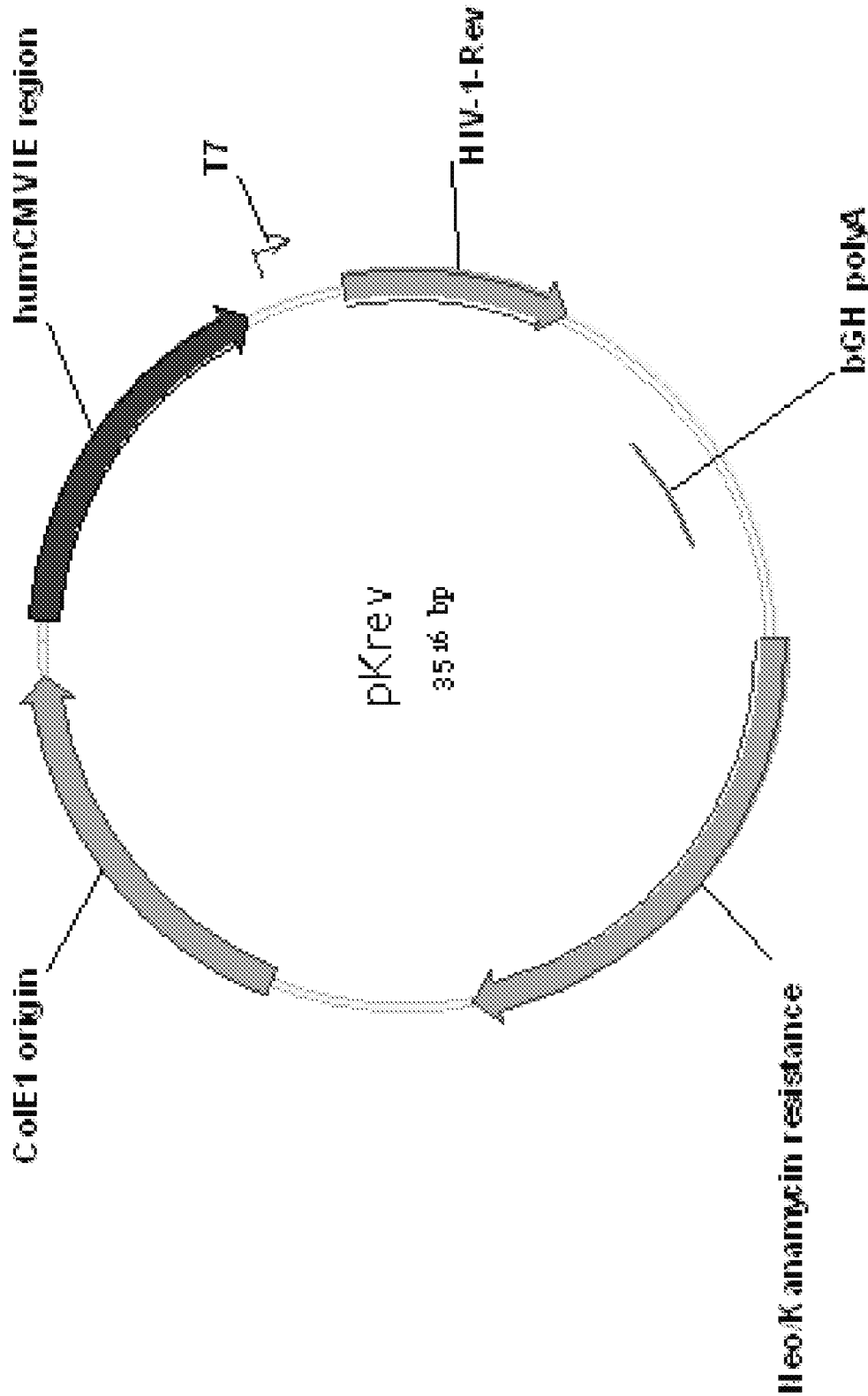


Figure 40.

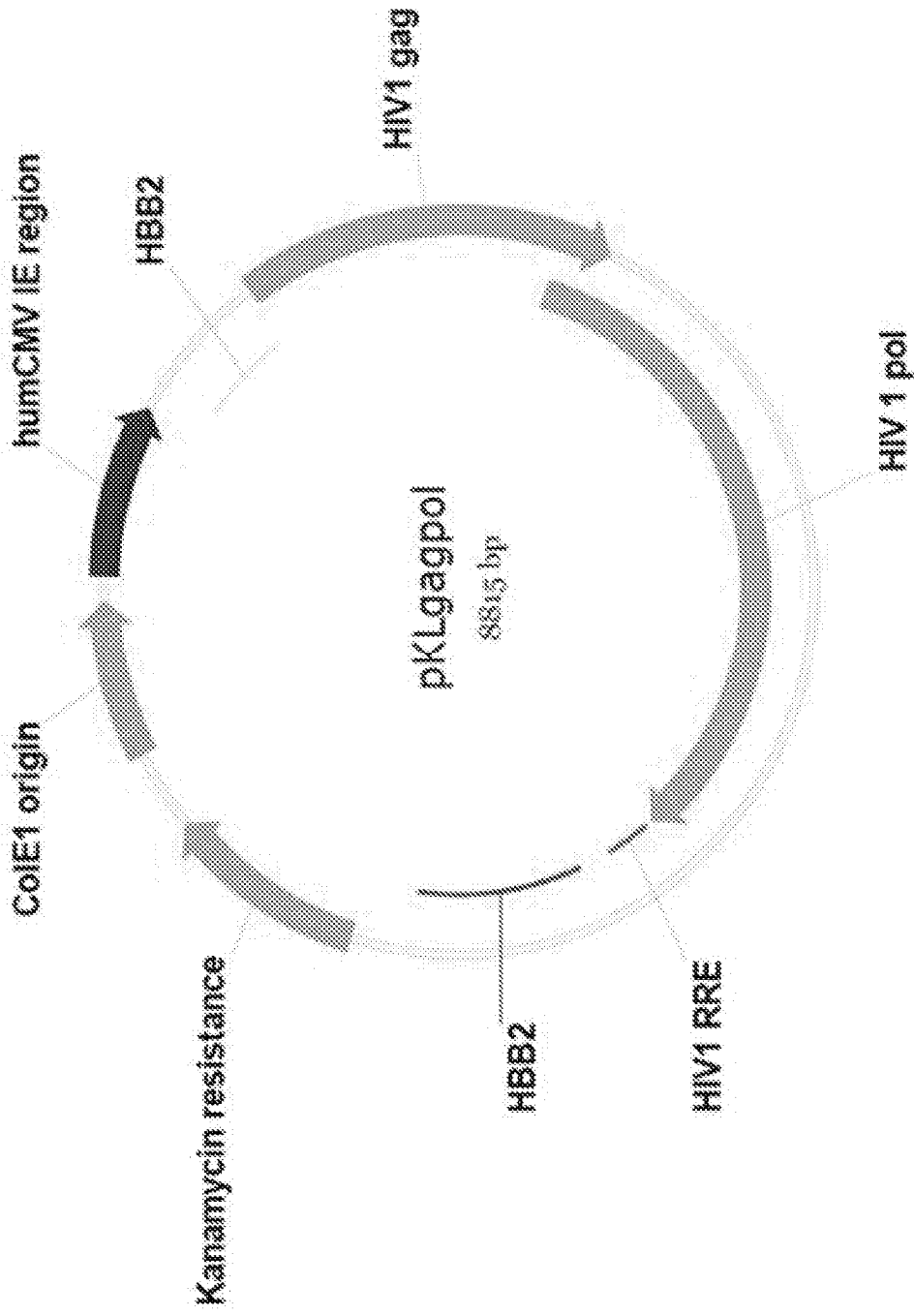


Figure 41.

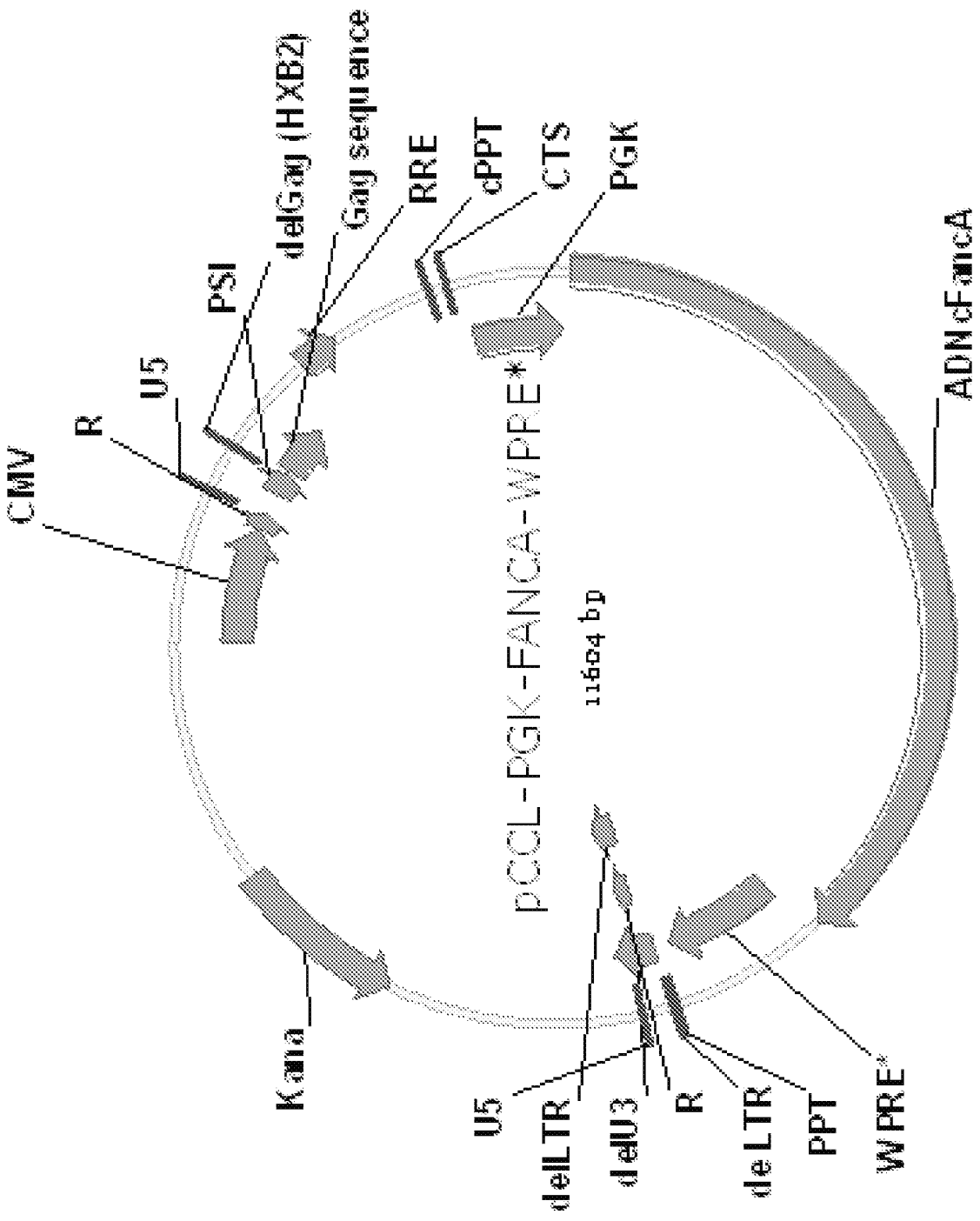


Figure 42.

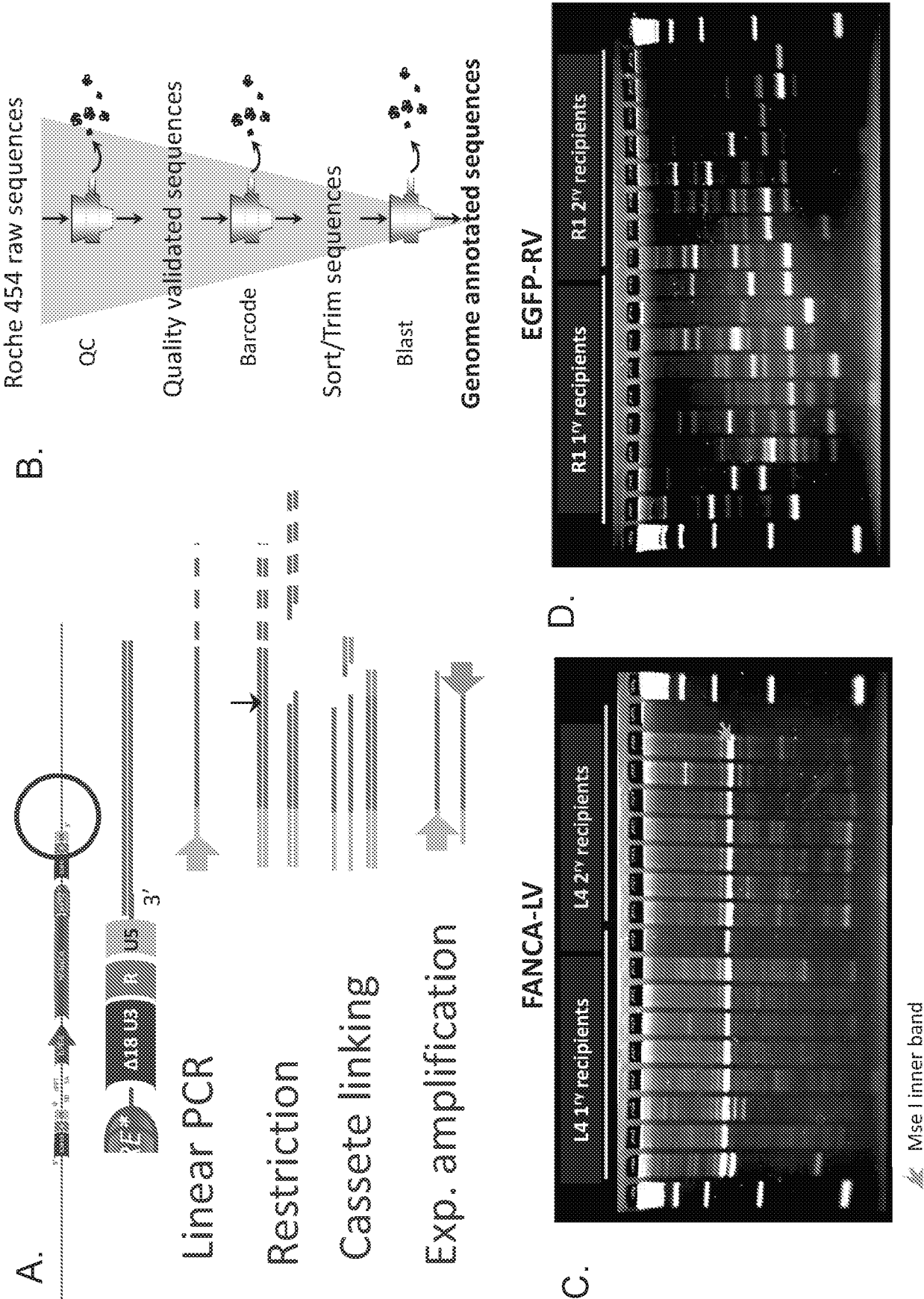
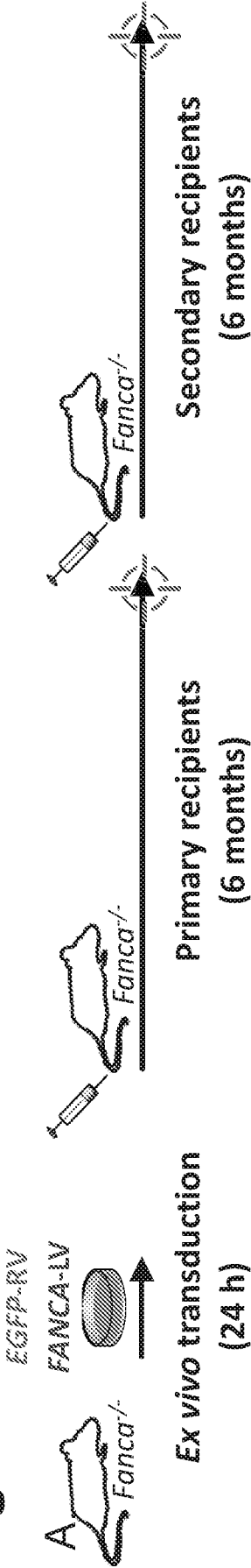


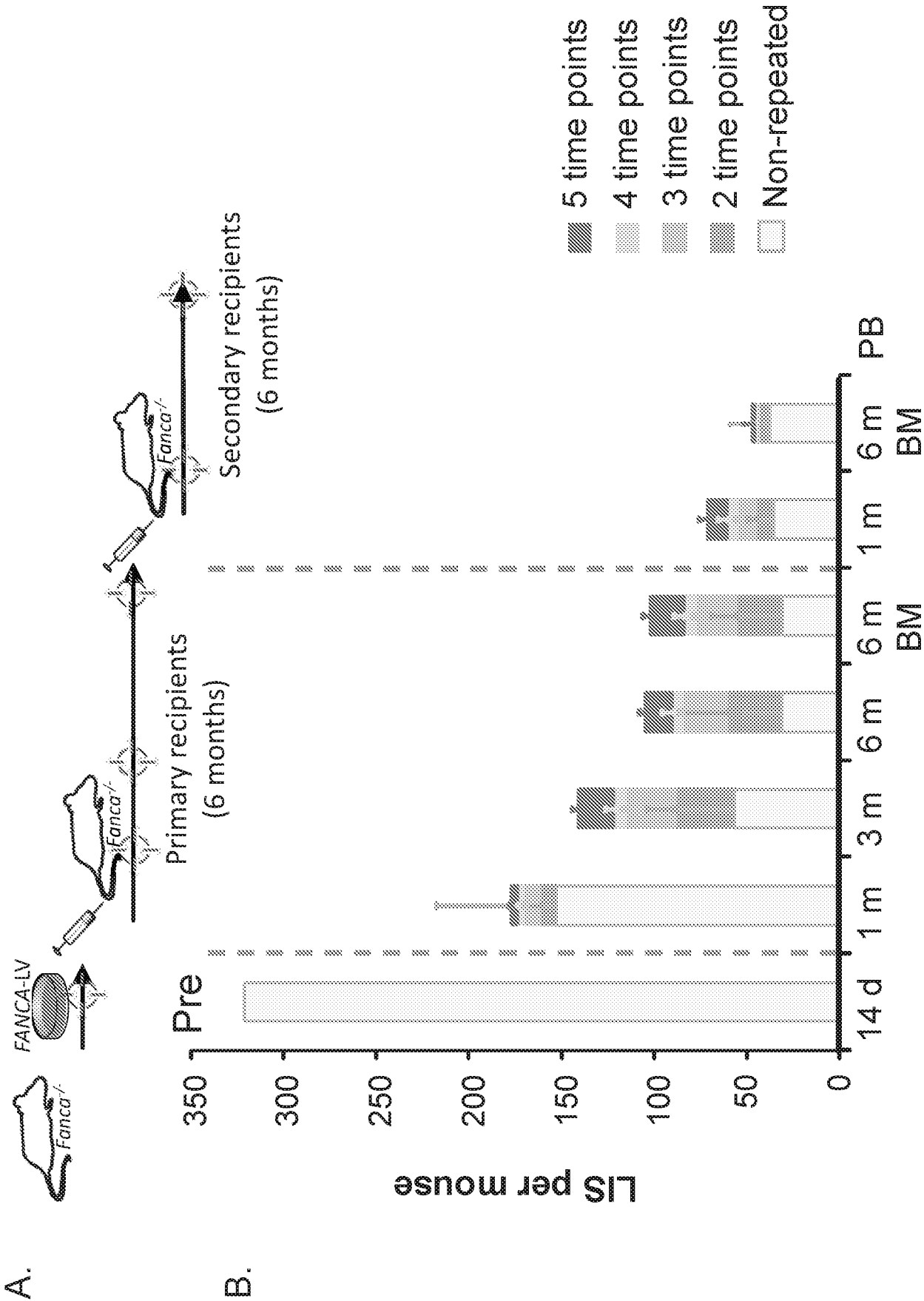
Figure 43



B.

	Mappable reads	Unique IS	RefSeq	RefSeq ±5 Kb	TSS ±5 Kb	CpG ±5 Kb
FANCA-LV						
Pre-BMT	1,646	321	199 62.0%	207 64.5%	19 5.9%	20 6.2%
Primary (n=31)	156,245	5,451	2915 53.5%	3129 57.4%	311 5.7%	359 6.6%
Secondary (n=25)	76,532	1,019	489 48.0%	526 51.6%	51 5.0%	58 5.7%
EGFP-RV						
Primary (n=9)	4,216	100	31 31.0%	40 40.0%	36 36.0%	13 16.3%
Secondary (n=8)	6,744	36	11 30.6%	14 38.9%	10 27.8%	3 12.3%

Figure 44.



SEQUENCE LISTING

<110> Centro de Investigaciones Energeticas Medioambientales y
Tecnologicas
Fundacion Instituto de Investigacion Sanitaria Fundacion
Jimenez Diaz
Centro de Investigacion Biomedica en Red
Fundacion para la investigacion biomedica Hospital Infantil
Universitario Nino Jesus
Bueren, Juan A.
Rio, Paula
Navarro, Susana
Sevilla, Julian
Segovia, Jose Carlos
Gonzalez, Africa
Casado, Jose Antonio
Guenechea, Guillermo

<120> GENE THERAPY FOR PATIENTS WITH FANCONI ANEMIA

<130> ROPA-002/01WO 329592-2007

<160> 26

<170> PatentIn version 3.5

<210> 1

<211> 234

<212> DNA

<213> Human immunodeficiency virus 1 (HIV-1)

<400> 1

aggagctttg ttccttgggt tcttgggagc agcaggaagc actatgggcg cagcgtcaat 60

gacgctgacg gtacaggcca gacaattatt gtctggtata gtgcagcagc agaacaattt 120

gctgagggct attgaggcgc aacagcatct gttgcaactc acagtctggg gcatcaagca 180

gctccaggca agaatcctgg ctgtggaaaag atacctaaag gatcaacagc tcct 234

<210> 2

<211> 126

<212> DNA

<213> Human immunodeficiency virus 1 (HIV-1)

<400> 2

ctctctcgac gcaggactcg gcttgctgaa gcgcgcacgg caagaggcga ggggcggcga 60

10 Aug 2021 2017322511

ctggtgagta cgccaaaaat tttgactagc ggaggctaga aggagagaga tgggtgcgag	120
agcgtc	126
<210> 3	
<211> 181	
<212> DNA	
<213> Human immunodeficiency virus 1 (HIV-1)	
<400> 3	
gggtctctct ggtagacca gatctgagcc tgggagctct ctggctaact agggaaccca	60
ctgcttaagc ctcaataaag cttgccttga gtgcttcaag tagtgtgtgc ccgtctgttg	120
tgtgactctg gtaactagag atccctcaga cccttttagt cagtgtggaa aatctctagc	180
a	181
<210> 4	
<211> 234	
<212> DNA	
<213> Human immunodeficiency virus 1 (HIV-1)	
<400> 4	
tggaagggtc aattcactcc caacgaagac aagatctgct ttttgcttgt actgggtctc	60
tctggttaga ccagatctga gcctgggagc tctctggcta actaggaac cactgctta	120
agcctcaata aagcttgctt tgagtgttc aagtagtgtg tgcccgctct ttgtgtgact	180
ctggtaacta gagatccctc agaccctttt agtcagtgtg gaaaatctct agca	234
<210> 5	
<211> 204	
<212> DNA	
<213> Human herpesvirus-5	
<400> 5	
gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt	60
ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac	120
tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg	180
tgggaggtct atataagcag agct	204

```

<210> 6
<211> 118
<212> DNA
<213> Human immunodeficiency virus 1 (HIV-1)

<400> 6
ttttaaaaga aaagggggga ttggggggta cagtgcaggg gaaagaatag tagacataat      60
agcaacagac atacaaaacta aagaattaca aaaacaaatt acaaaaaattc aaaatttt      118

<210> 7
<211> 511
<212> DNA
<213> Homo sapiens

<400> 7
ggggttgggg ttgcgccttt tccaaggcag ccctggggtt gcgcagggac gcggctgctc      60
tgggcgtggt tccgggaaac gcagcggcgc cgaccctggg tctcgacat tcttcacgtc      120
cgttcgcagc gtcacccgga tcttcgccgc tacccttggt ggccccccgg cgacgcttcc      180
tgctccgccc ctaagtcggg aaggttcctt gcggttcgcg gcgtgccgga cgtgacaaac      240
ggaagccgca cgtctcacta gtaccctcgc agacggacag cgccaggagg caatggcagc      300
gcgccgaccg cgatgggctg tggccaatag cggctgctca gcagggcgcg ccgagagcag      360
cggccgggaa ggggcgggtgc gggaggcggg gtgtggggcg gtagtgtggg ccctgttcct      420
gcccgcgcgg tgttccgcat tctgcaagcc tccggagcgc acgtcggcag tcggctccct      480
cgttgaccga atcaccgacc tctctcccca g                                511

<210> 8
<211> 4368
<212> DNA
<213> Homo sapiens

<400> 8
atgtccgact cgtgggtccc gaactccgcc tcgggccagg acccaggggg ccgccggagg      60
gcctgggccc agctgctggc gggaagggtc aagagggaat aatataatcc tgaaagggca      120
cagaaattaa aggaatcagc tgtgcgcctc ctgcgaagcc atcaggacct gaatgccctt      180

```

ttgcttgagg tagaagggtcc actgtgtaaa aaattgtctc tcagcaaagt gattgactgt	240
gacagttctg aggcctatgc taatcattct agttcattta taggctctgc tttgcaggat	300
caagcctcaa ggctgggggt tcccgtgggt attctctcag ccgggatggg tgcctctagc	360
gtgggacaga tctgcacggc tccagcggag accagtcacc ctgtgctgct gactgtggag	420
cagagaaaaga agctgtcttc cctgttagag tttgctcagt atttattggc acacagtatg	480
ttctcccgtc tttccttctg tcaagaatta tggaaaatac agagttcttt gttgcttgaa	540
gcgggtgtggc atcttcacgt acaaggcatt gtgagcctgc aagagctgct ggaaagccat	600
cccgacatgc atgctgtggg atcgtggctc ttcaggaatc tgtgctgcct ttgtgaacag	660
atggaagcat cctgccagca tgctgacgtc gccagggccca tgctttctga ttttgttcaa	720
atgtttgttt tgaggggatt tcagaaaaac tcagatctga gaagaactgt ggagcctgaa	780
aaaatgccgc aggtcacggg tgatgtactg cagagaatgc tgatttttgc acttgacgct	840
ttggctgctg gagtacagga ggagtcctcc actcacaaga tcgtgaggtg ctggttcgga	900
gtgttcagtg gacacacgct tggcagtgtg atttccacag atcctctgaa gaggttcttc	960
agtcataccc tgactcagat actcactcac agccctgtgc tgaaagcatc tgatgctgtt	1020
cagatgcaga gagagtggag ctttgcgagg acacaccctc tgctcacctc actgtaccgc	1080
aggctctttg tgatgctgag tgcagaggag ttggttggcc atttgcaaga agttctggaa	1140
acgcaggagg ttacttgcca gagagtgtc tcctttgtgt ctgccctggg tgtctgcttt	1200
ccagaagcgc agcagctgct tgaagactgg gtggcgcggt tgatggccca ggcattcgag	1260
agctgccagc tggacagcat ggtcactgcg ttcctgggtg tgcgccaggc agcactggag	1320
ggccccctctg cgttcctgtc atatgcagac tggttcaagg cctcctttgg gagcacacga	1380
ggctaccatg gctgcagcaa gaaggccctg gtcttcctgt ttacgttctt gtcagaactc	1440
gtgccttttg agtctccccg gtacctgcag gtgcacattc tccaccacc cctggttccc	1500
agcaagtacc gctccctcct cacagactac atctcattgg ccaagacacg gctggccgac	1560
ctcaaggttt ctatagaaaa catgggactc tacgaggatt tgtcatcagc tggggacatt	1620
actgagcccc acagccaagc tcttcaggat gttgaaaagg ccatcatggt gtttgagcat	1680

acggggaaca	tcccagtcac	cgatcatggag	gccagcatat	tcaggaggcc	ttactacgtg	1740
tcccacttcc	tccccgccct	gctcacacct	cgagtgtctcc	ccaaagtccc	tgactcccgt	1800
gtggcgttta	tagagtctct	gaagagagca	gataaaatcc	ccccatctct	gtactccacc	1860
tactgccagg	cctgctctgc	tgctgaagag	aagccagaag	atgcagccct	gggagtgagg	1920
gcagaaccca	actctgctga	ggagcccctg	ggacagctca	cagctgcact	gggagagctg	1980
agagcctcca	tgacagaccc	cagccagcgt	gatgttatat	cggcacaggt	ggcagtgatt	2040
tctgaaagac	tgagggctgt	cctgggccac	aatgaggatg	acagcagcgt	tgagatatca	2100
aagattcagc	tcagcatcaa	cacgccgaga	ctggagccac	gggaacacat	tgctgtggac	2160
ctcctgctga	cgtctttctg	tcagaacctg	atggctgcct	ccagtgtcgc	tcccccgag	2220
aggcagggtc	cctgggctgc	cctcttcgtg	aggaccatgt	gtggacgtgt	gctccctgca	2280
gtgctcacc	ggctctgcca	gctgctccgt	caccagggcc	cgagcctgag	tgccccacat	2340
gtgctgggg	tggctgccct	ggccgtgcac	ctgggtgagt	ccaggctctg	gctcccagag	2400
gtggatgtgg	gtcctcctgc	acctgggtg	ggccttcctg	tccctgcgct	ctttgacagc	2460
ctcctgacct	gtaggacgag	ggattccttg	ttcttctgcc	tgaaattttg	tacagcagca	2520
atttcttact	ctctctgcaa	gttttcttcc	cagtcacgag	atactttgtg	cagctgctta	2580
tctccaggcc	ttattaaaaa	gtttcagttc	ctcatgttca	gattgttctc	agaggcccga	2640
cagcctcttt	ctgaggagga	cgtagccagc	ctttcctgga	gacccttgca	ccttccttct	2700
gcagactggc	agagagctgc	cctctctctc	tggacacaca	gaaccttccg	agagggtgtg	2760
aaagaggaag	atgttcactt	aacttaccaa	gactgggttac	acctggagct	ggaaattcaa	2820
cctgaagctg	atgctctttc	agatactgaa	cggcaggact	tccaccagt	ggcgatccat	2880
gagcactttc	tccctgagtc	ctcggcttca	gggggctgtg	acggagacct	gcaggctgcg	2940
tgtaccattc	ttgtcaacgc	actgatggat	ttccaccaa	gctcaaggag	ttatgaccac	3000
tcagaaaatt	ctgatttggt	ctttgggtggc	cgcacaggaa	atgaggatat	tatttccaga	3060
ttgcaggaga	tggtagctga	cctggagctg	cagcaagacc	tcatagtgcc	tctcggccac	3120

```

accccttccc aggagcactt cctctttgag attttccgca gacggctcca ggctctgaca      3180
agcgggtgga gcgtggctgc cagccttcag agacagaggg agctgctaata gtacaaacgg      3240
atcctcctcc gcctgccttc gtctgtcctc tgcggcagca gcttccaggc agaacagccc      3300
atcactgcca gatgcgagca gttcttccac ttggtcaact ctgagatgag aaacttctgc      3360
ttccacggag gtgccctgac acaggacatc actgcccact tcttcagggg cctcctgaac      3420
gcctgtctgc ggagcagaga cccctccctg atggctcgact tcatactggc caagtgccag      3480
acgaaatgcc ccttaatttt gacctctgct ctgggtgtggg ggccgagcct ggagcctgtg      3540
ctgctctgcc ggtggaggag aactgccag agcccgctgc cccgggaact gcagaagcta      3600
caagaaggcc ggagtttgc cagcgatttc ctctcccctg aggctgcctc cccagcacc      3660
aaccggact ggctctcagc tgctgcactg cactttgcga ttcaacaagt cagggaagaa      3720
aacatcagga agcagctaaa gaagctggac tgcgagagag aggagctatt ggttttcctt      3780
ttcttcttct ccttgatggg cctgctgtcg tcacatctga cctcaaatag caccacagac      3840
ctgcaaaagg ctttccacgt ttgtgcagca atcctcgagt gtttagagaa gaggaagata      3900
tcctggctgg cactctttca gttgacagag agtgacctca ggctggggcg gctcctcctc      3960
cgtgtggccc cggatcagca caccaggctg ctgcctttcg ctttttacag tcttctctcc      4020
tacttccatg aagacgcggc catcagggaag gaggccttcc tgcattgtgc tgtggacatg      4080
tacttgaagc tgggtccagct cttcgtggct ggggatacaa gcacagtttc acctccagct      4140
ggcaggagcc tggagctcaa gggtcagggc aaccccgctg aactgataac aaaagctcgt      4200
ctttttctgc tgcagttaat acctcggtgc ccgaaaaaga gcttctcaca cgtggcagag      4260
ctgctggctg atcgtgggga ctgcgaccca gaggtgagcg ccgccctcca gagcagacag      4320
caggctgccc ctgacgctga cctgtcccag gagcctcatc tcttctga      4368

```

```

<210> 9
<211> 380
<212> DNA
<213> Human herpesvirus-5

<400> 9

```

```

gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc      60
catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgcca      120
acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga      180
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc      240
aagtgtatca tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgctt      300
ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat      360
tagtcatcgc tattaccatg      380

<210> 10
<211> 122
<212> DNA
<213> Simian virus 40

<400> 10
aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca      60
aataaagcat ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct      120
ta      122

<210> 11
<211> 136
<212> DNA
<213> Simian virus 40

<400> 11
atcccccccc taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt      60
tttatttatg cagaggccga ggccgcctcg gcctctgagc tattccagaa gtagtgagga      120
ggcttttttg gaggcc      136

<210> 12
<211> 28
<212> DNA
<213> Human immunodeficiency virus 1 (HIV-1)

<400> 12
tttaaaagaa aaggggggat tgggggggt      28

```

```

<210> 13
<211> 83
<212> DNA
<213> Human immunodeficiency virus 1 (HIV-1)

<400> 13
gaattcgagc tcggtacctt taagaccaat gacttacaag gcagctgtag atcttagcca      60

ctttttaaaa gaaaaggggg gac                                             83

<210> 14
<211> 795
<212> DNA
<213> Escherichia coli

<400> 14
atgattgaac aagatggatt gcacgcaggt tctccggcgg cttgggtgga gaggctattc      60

ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca      120

gcgcgaggggc gtccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg      180

caagacgagg cagcgcggtc atcgtggctg gcgacgacgg gcgttccttg cgcggtgtg      240

ctcgacgttg tcaactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag      300

gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg      360

cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc      420

atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa      480

gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgtc tatgcccgcac      540

ggcgaggatc tcgtcgtgac ccacggcgat gcctgcttgc cgaatatcat ggtggaaaaat      600

ggccgctttt ctggattcat cgactgtggc cgtctgggtg tggcggaccg ctatcaggac      660

atagcgttgg ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc      720

cttgtgcttt acggtatcgc cgcgcccgat tcgcagcgca tcgccttcta tcgccttctt      780

gacgagttct tctga                                                         795

<210> 15
<211> 137

```

```

<212> DNA
<213> Escherichia coli

<400> 15
gcattggcgc agaaaaaat gcctgatgcg acgctgcgcg tcttatactc ccacatatgc      60
cagattcagc aacggatacg gcttcccca cttgcccact tccatacgtg tcctccttac      120
cagaaattta tccttaa                                           137

```

```

<210> 16
<211> 589
<212> DNA
<213> unknown

<220>
<223> ColE1/pMB1/pBR322/pUC origin of replication plasmid sequence

<400> 16
ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc      60
agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt      120
cagcagagcg cagataccaa atactgttct tctagtgtag ccgtagttag gccaccactt      180
caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc      240
tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa      300
ggcgcagcgg tcgggctgaa cgggggggtt gtgcacacag cccagcttgg agcgaacgac      360
ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc ttcccgaagg      420
gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc gcacgaggga      480
gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact      540
tgagcgtcga tttttgtgat gtcgctcagg ggggcggagc ctatggaaa           589

```

```

<210> 17
<211> 22
<212> DNA
<213> Escherichia coli

<400> 17
taatgtgagt tagctcactc at                                           22

```

10 Aug 2021

2017322511

<210> 18
 <211> 31
 <212> DNA
 <213> Escherichia coli

 <400> 18
 tttaacacattt atgcttccgg ctcgtatggt g 31

<210> 19
 <211> 17
 <212> DNA
 <213> Escherichia coli

 <400> 19
 ttgtgagcgg ataacaa 17

<210> 20
 <211> 19
 <212> DNA
 <213> Bacteriophage T3

 <400> 20
 aattaaccct cactaaagg 19

<210> 21
 <211> 19
 <212> DNA
 <213> Bacteriophage T7

 <400> 21
 cctatagtga gtcgtatta 19

<210> 22
 <211> 428
 <212> DNA
 <213> F1 bacteriophage

 <400> 22
 acgcgccctg tagcggcgca ttaagcgcgg cgggtgtggt ggttacgcgc agcgtgaccg 60
 ctacacttgc cagcggcccta gcgcccgtc ctttcgcttt cttcccttcc tttctcgcca 120
 cgttcgccgg ctttccccgt caagctctaa atcggggggt ccctttagggt ttccgattta 180

```

gtgcttttacg gcacctcgac cccaaaaaac ttgattaggg tgatgggttca cgtagtgggc      240
catcgccctg atagacgggt tttcgccctt tgacgttggg gtccacgttc tttaatagtg      300
gactcttggt ccaaactgga acaacactca accctatctc ggtctattct tttgatttat      360
aagggatttt gccgatttcg gcctattggg taaaaaatga gctgatttaa caaaaattta      420
acgcgaat                                         428

```

```

<210> 23
<211> 677
<212> DNA
<213> Woodchuck hepatitis virus

```

```

<400> 23
cgagcatctt accgccattt attcccatat ttgttctggt tttcttgatt tgggtataca      60
tttaaagtgt aataaaacaa aatgggtgggg caatcattta catttttagg gatatgtaat      120
tactagttca ggtgtattgc cacaagacaa acatgttaag aaactttccc gttattttacg      180
ctctgttcct gttaatcaac ctctggatta caaaatttgt gaaagattga ctgatattct      240
taactatgtt gtccttttta cgctgtgtgg atatgctgct ttaatgcctc tgtatcatgc      300
tattgcttcc cgtacggctt tcgttttctc ctcttgtat aaatcctggg tgctgtctct      360
ttatgaggag ttgtggcccg ttgtccgtca acgtggcgtg gtgtgctctg tgtttgctga      420
cgcaaccccc actggctggg gcattgccac cacctgtcaa ctcttttctg ggactttcgc      480
tttccccctc ccgatcgcca cggcagaact catcgccgcc tgccttgccc gctgctggac      540
aggggctagg ttgctgggca ctgataattc cgtgggtgtg tcggggaagg gcctgctgcc      600
ggctctgcgg cctcttccgc gtcttcgcct tcgccctcag acgagtcgga tctcccttg      660
ggccgcctcc ccgcctg                                         677

```

```

<210> 24
<211> 11433
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Made in lab transfer cassette pCCL-SIN-cPPT/CTS -hPGK-hFANCA-WPRE

```

```

<400> 24
catgacaaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 60
gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaacaaa 120
aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa gagctaccaa ctctttttcc 180
gaaggtaact ggcttcagca gagcgcagat accaaatact gttcttctag tgtagccgta 240
gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaatacct 300
gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg 360
atagttaccg gataaggcgc agcggtcggg ctgaacgggg gggttcgtgca cacagcccag 420
cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaaagcgc 480
cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg tcggaacagg 540
agagcgcacg agggagcttc caggggggaaa cgcttggtat ctttatagtc ctgtcggggt 600
tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg 660
gaaaaacgcc agcaacgcgg cttttttacg gttcctggcc ttttgctggc cttttgctca 720
catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cttttgagtg 780
agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc 840
ggaagagcgc ccaatacgca aaccgcctct ccccgcgctg tggccgattc attaatacag 900
ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa ttaatgtgag 960
ttagctcact cattaggcac cccaggcttt acactttatg cttccggctc gtatgttgtg 1020
tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg attacgcaa 1080
gcgcgcaatt aaccctcact aaagggaaca aaagctggag ctgcaagctt ggccattgca 1140
tacgttgatc ccatatcata atatgtacat ttatattggc tcatgtccaa cattaccgcc 1200
atgttgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 1260
tagcccatat atggagttcc gcgttacata acttacggta aatggcccg ctaggctgacc 1320
gccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 1380
agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 1440

```

acatcaagtg tatcatatgc caagtacgcc ccctattgac gtcaatgacg gtaaatggcc	1500
cgcctggcat tatgcccagt acatgacctt atgggacttt cctacttggc agtacatcta	1560
cgtattagtc atcgctatta ccatgggtgat gcggtttttg cagtacatca atgggcgtgg	1620
atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt	1680
gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactccg cccattgac	1740
gcaaatgggc ggtaggcgtg tacgggtggga ggtctatata agcagagctc gtttagtgaa	1800
ccgggggtctc tctggttaga ccagatctga gcctggggagc tctctggcta actagggaac	1860
ccactgctta agcctcaata aagcttgctt tgagtgcctt aagtagtgtg tgcccgtctg	1920
tttgttgact ctggtaacta gagatccctc agaccctttt agtcagtgtg gaaaatctct	1980
agcagtggcg cccgaacagg gacttgaaag cgaaagggaa accagaggag ctctctcgac	2040
gcaggactcg gcttgctgaa gcgcgcacgg caagaggcga ggggcggcga ctggtgagta	2100
cgccaaaaat tttgactagc ggaggctaga aggagagaga tgggtgcgag agcgtcagta	2160
ttaagcgggg gagaattaga tcgcgatggg aaaaaattcg gttaaggcca gggggaaaga	2220
aaaaatataa attaaaacat atagtatggg caagcaggga gctagaacga ttcgcagtta	2280
atcctggcct gttagaaaca tcagaaggct gtagacaaat actgggacag ctacaaccat	2340
cccttcagac aggatcagaa gaacttagat cattatataa tacagtagca accctctatt	2400
gtgtgcatca aaggatagag ataaaagaca ccaaggaagc tttagacaag atagaggaag	2460
agcaaaacaa aagtaagacc accgcacagc aagcggccgc tgatcttcag acctggagga	2520
ggagatatga gggacaattg gagaagtga ttatataaat ataaagtagt aaaaattgaa	2580
ccattaggag tagcaccac caaggcaaag agaagagtgg tgcagagaga aaaaagagca	2640
gtgggaatag gagctttgtt ccttgggttc ttgggagcag caggaagcac tatgggcgca	2700
gcgtcaatga cgctgacggt acaggccaga caattattgt ctggtatagt gcagcagcag	2760
aacaatttgc tgagggtat tgaggcgcaa cagcatctgt tgcaactcac agtctggggc	2820
atcaagcagc tccaggcaag aatcctggct gtggaaagat acctaaagga tcaacagctc	2880

ctggggattt	gggggttgctc	tggaaaactc	atttgcacca	ctgctgtgcc	ttggaatgct	2940
agttggagta	ataaatctct	ggaacagatt	tggaatcaca	cgacctggat	ggagtgggac	3000
agagaaatta	acaattacac	aagcttaata	cactccttaa	ttgaagaatc	gcaaaaccag	3060
caagaaaaga	atgaacaaga	attattggaa	ttagataaat	gggcaagttt	gtggaattgg	3120
tttaacataa	caaattggct	gtggtatata	aaattattca	taatgatagt	aggaggcttg	3180
gtaggtttta	gaatagtttt	tgctgtactt	tctatagtga	atagagttag	gcagggatat	3240
tcaccattat	cgtttcagac	ccacctccca	accccgaggg	gacccgacag	gcccgaagga	3300
atagaagaag	aagggtggaga	gagagacaga	gacagatcca	ttcgattagt	gaacggatct	3360
cgacggtatc	ggttaacttt	taaaagaaaa	ggggggattg	gggggtacag	tgacggggaa	3420
agaatagtag	acataatagc	aacagacata	caaaactaaag	aattacaaaa	acaaattaca	3480
aaaattcaaa	attttatcga	tcacgagact	agcctcgaga	agcttgatat	cgaattccac	3540
gggggttggg	ttgcgccttt	tccaaggcag	ccctgggttt	gcgcagggac	gcggctgctc	3600
tgggcgtggt	tccgggaaac	gcagcggcgc	cgaccctggg	tctcgcacat	tcttcacgtc	3660
cgttcgcagc	gtcaccgga	tcttcgccgc	tacccttggt	ggcccccg	cgacgcttcc	3720
tgctccgccc	ctaagtcggg	aaggttcctt	gcggttcgcg	gcgtgccgga	cgtgacaaac	3780
ggaagccgca	cgtctcacta	gtaccctcgc	agacggacag	cgccaggag	caatggcagc	3840
gcgccgaccg	cgatgggctg	tggccaatag	cggctgctca	gcagggcgcg	ccgagagcag	3900
cggccgggaa	ggggcggtgc	gggaggcggg	gtgtggggcg	gtagtgtggg	ccctgttcct	3960
gcccgcgcgg	tgttccgcat	tctgcaagcc	tccggagcgc	acgtcggcag	tcggctccct	4020
cgttgaccga	atcaccgacc	tctctcccca	gggggatccc	ccgggctgca	ggaattcatg	4080
tccgactcgt	gggtcccga	ctccgcctcg	ggccaggacc	cagggggccg	ccggagggcc	4140
tgggccgagc	tgctggcggg	aagggtcaag	agggaaaaat	ataatcctga	aagggcacag	4200
aaattaaagg	aatcagctgt	gcgcctcctg	cgaagccatc	aggacctgaa	tgcccttttg	4260
cttgaggtag	aaggctccact	gtgtaaaaaa	ttgtctctca	gcaaagtgat	tgactgtgac	4320
agttctgagg	cctatgctaa	tcattctagt	tcatttatag	gctctgcttt	gcaggatcaa	4380

gcctcaaggc tgggggttcc cgtgggtatt ctctcagccg ggatggttgc ctctagcgtg	4440
ggacagatct gcacggctcc agcggagacc agtcaccctg tgctgctgac tgtggagcag	4500
agaaagaagc tgtcttccct gttagagttt gctcagtatt tattggcaca cagtatgttc	4560
tcccgctctt ccttctgtca agaattatgg aaaatacaga gttctttgtt gcttgaagcg	4620
gtgtggcatc ttcacgtaca aggcatgtg agcctgcaag agctgctgga aagccatccc	4680
gacatgcatg ctgtgggatc gtggctcttc aggaatctgt gctgcctttg tgaacagatg	4740
gaagcatcct gccagcatgc tgacgtcgcc agggccatgc tttctgattt tgttcaaatg	4800
tttgttttga ggggatttca gaaaaactca gatctgagaa gaactgtgga gcctgaaaaa	4860
atgccgcagg tcacggttga tgtactgcag agaatgctga tttttgcact tgacgctttg	4920
gctgctggag tacaggagga gtcctccact cacaagatcg tgagggtgctg gttcggagtg	4980
ttcagtggac acacgcttgg cagtgttaatt tccacagatc ctctgaagag gttcttcagt	5040
cataccctga ctcagatact cactcacagc cctgtgctga aagcatctga tgctgttcag	5100
atgcagagag agtggagctt tgcgcggaca caccctctgc tcacctcact gtaccgcagg	5160
ctctttgtga tgctgagtg agaggagtgt gttggccatt tgcaagaagt tctggaaacg	5220
caggagggtc actggcagag agtgctctcc tttgtgtctg ccctggttgt ctgctttcca	5280
gaagcgcagc agctgcttga agactgggtg gcgcgtttga tggcccaggc attcgagagc	5340
tgccagctgg acagcatggt cactgcgttc ctggttgtgc gccaggcagc actggagggc	5400
ccctctgcgt tcctgtcata tgcagactgg ttcaaggcct cttttgggag cacacgaggc	5460
taccatggct gcagcaagaa ggccctggtc ttcctgttta cgttcttgtc agaactcgtg	5520
ccttttgagt ctccccggtc cctgcagggtg cacattctcc acccaccctt ggttcccagc	5580
aagtaccgct ccctcctcac agactacatc tcattggcca agacacggct ggccgacctc	5640
aaggtttcta tagaaaacat gggactctac gaggatttgt catcagctgg ggacattact	5700
gagccccaca gccaagctct tcaggatgtt gaaaaggcca tcatggtgtt tgagcatacg	5760
gggaacatcc cagtcaccgt catggaggcc agcatattca ggaggcctta ctacgtgtcc	5820

cacttcctcc	ccgccctgct	cacacctcga	gtgctcccca	aagtcctga	ctcccgtgtg	5880
gcgtttatag	agtctctgaa	gagagcagat	aaaaatcccc	catctctgta	ctccacctac	5940
tgccaggcct	gctctgctgc	tgaagagaag	ccagaagatg	cagccctggg	agtgagggca	6000
gaacccaact	ctgctgagga	gcccctggga	cagctcacag	ctgcactggg	agagctgaga	6060
gcctccatga	cagaccccag	ccagcgtgat	gttatatcgg	cacaggtggc	agtgatttct	6120
gaaagactga	gggctgtcct	gggccacaat	gaggatgaca	gcagcgttga	gatatcaaag	6180
attcagctca	gcatcaacac	gccgagactg	gagccacggg	aacacattgc	tgtggacctc	6240
ctgctgacgt	ctttctgtca	gaacctgatg	gctgcctcca	gtgtcgctcc	cccggagagg	6300
cagggtccct	gggctgccct	cttcgtgagg	accatgtgtg	gacgtgtgct	ccctgcagtg	6360
ctcaccggc	tctgccagct	gctccgtcac	caggggcccga	gcctgagtgc	cccacatgtg	6420
ctgggggttg	ctgccctggc	cgtgcacctg	ggtgagtcca	ggctctcgct	cccagagggtg	6480
gatgtgggtc	ctcctgcacc	tggtgctggc	cttcctgtcc	ctgcgctctt	tgacagcctc	6540
ctgacctgta	ggacgaggga	ttccttgttc	ttctgcctga	aattttgtac	agcagcaatt	6600
tcttactctc	tctgcaagtt	ttcttcccag	tcacgagata	ctttgtgcag	ctgcttatct	6660
ccaggcctta	ttaaaaagtt	tcagttcctc	atgttcagat	tgttctcaga	ggcccgacag	6720
cctctttctg	aggaggacgt	agccagcctt	tcctggagac	ccttgcacct	tccttctgca	6780
gactggcaga	gagctgccct	ctctctctgg	acacacagaa	ccttccgaga	ggtgttgaaa	6840
gaggaagatg	ttcacttaac	ttaccaagac	tggttacacc	tggagctgga	aattcaacct	6900
gaagctgatg	ctctttcaga	tactgaacgg	caggacttcc	accagtgggc	gatccatgag	6960
cactttctcc	ctgagtcctc	ggcttcaggg	ggctgtgacg	gagacctgca	ggctgcgtgt	7020
accattcttg	tcaacgcact	gatggatttc	caccaaagct	caaggagtta	tgaccactca	7080
gaaaattctg	atttggctct	tggtggccgc	acaggaaatg	aggatattat	ttccagattg	7140
caggagatgg	tagctgacct	ggagctgcag	caagacctca	tagtgcctct	cggccacacc	7200
ccttcccagg	agcacttcct	ctttgagatt	ttccgcagac	ggctccaggc	tctgacaagc	7260
gggtggagcg	tggctgccag	ccttcagaga	cagagggagc	tgctaattga	caaacggatc	7320

ctcctccgcc	tgcttctgtc	tgtcctctgc	ggcagcagct	tccaggcaga	acagcccatc	7380
actgccagat	gcgagcagtt	cttccacttg	gtcaactctg	agatgagaaa	cttctgtctc	7440
cacggagggtg	ccctgacaca	ggacatcact	gcccacttct	tcaggggcct	cctgaacgcc	7500
tgtctgcgga	gcagagaccc	ctccctgatg	gtcgacttca	tactggccaa	gtgccagacg	7560
aaatgcccct	taattttgac	ctctgctctg	gtgtggtggc	cgagcctgga	gcctgtgctg	7620
ctctgccggt	ggaggagaca	ctgccagagc	ccgctgcccc	gggaactgca	gaagctacaa	7680
gaaggccggc	agtttgccag	cgattttctc	tcccctgagg	ctgcctcccc	agcacccaac	7740
ccggactggc	tctcagctgc	tgactgcac	tttgcgattc	aacaagtcag	ggaagaaaac	7800
atcaggaagc	agctaaagaa	gctggactgc	gagagagagg	agctattggt	tttccttttc	7860
ttctttctcct	tgatgggcct	gctgtcgtca	catctgacct	caaatagcac	cacagacctg	7920
ccaaaggctt	tccacgtttg	tgacagcaatc	ctcgagtgtt	tagagaagag	gaagatatcc	7980
tggctggcac	tctttcagtt	gacagagagt	gacctcaggc	tggggcggct	cctcctccgt	8040
gtggccccgg	atcagcacac	caggctgctg	cctttcgctt	tttacagtct	tctctcctac	8100
ttccatgaag	acgcggccat	cagggaagag	gccttcctgc	atgttgctgt	ggacatgtac	8160
ttgaagctgg	tccagctctt	cgtggctggg	gatacaagca	cagtttcacc	tccagctggc	8220
aggagcctgg	agctcaaggg	tcagggaac	cccgtggaac	tgataacaaa	agctcgtctt	8280
tttctgctgc	agttaatacc	tcggtgcccc	aaaaagagct	tctcacacgt	ggcagagctg	8340
ctggctgatc	gtggggactg	cgaccagag	gtgagcgccg	ccctccagag	cagacagcag	8400
gctgcccctg	acgctgacct	gtcccaggag	cctcatctct	tctgatgaga	attcgatatc	8460
aagcttatcg	ataccgtcga	atccccggg	ctgcaggaat	tcgagcatct	taccgccatt	8520
tattcccata	tttgttctgt	ttttcttgat	ttgggtatac	atttaaagt	taataaaaca	8580
aaatggtggg	gcaatcattt	acatttttag	ggatatgtaa	ttactagttc	aggtgtattg	8640
ccacaagaca	aacatgttaa	gaaactttcc	cgttatttac	gctctgttcc	tgttaatcaa	8700
cctctggatt	acaaaatttg	tgaaagattg	actgatattc	ttaactatgt	tgctcctttt	8760

acgctgtgtg gatatgctgc tttaatgcct ctgtatcatg ctattgcttc ccgtacggct	8820
ttcgttttct cctccttgta taaatcctgg ttgctgtctc tttatgagga gttgtggccc	8880
gttgtccgtc aacgtggcgt ggtgtgctct gtgtttgctg acgcaacccc cactggctgg	8940
ggcattgcca ccacctgtca actcctttct gggactttcg ctttccccct cccgatcgcc	9000
acggcagaac tcatcgccgc ctgccttgcc cgctgctgga caggggctag gttgctgggc	9060
actgataatt ccgtgggtgtt gtcggggaag ggcctgctgc cggctctgcg gcctcttccg	9120
cgtcttcgcc ttcgccctca gacgagtcgg atctcccttt gggccgcctc cccgcctgga	9180
attcgagctc ggtaccttta agaccaatga cttacaaggc agctgtagat cttagccact	9240
ttttaaaaga aaagggggga ctggaagggc taattcactc ccaacgaaga caagatctgc	9300
tttttgcttg tactgggtct ctctgggttag accagatctg agcctgggag ctctctggct	9360
aactagggaa cccactgctt aagcctcaat aaagcttgcc ttgagtgcct caagtagtgt	9420
gtgcccgtct gttgtgtgac tctggtaact agagatccct cagacccttt tagtcagtgt	9480
ggaaaatctc tagcagtagt agttcatgtc atcttattat tcagtattta taacttgcaa	9540
agaaatgaat atcagagagt gagagggaact tgtttattgc agcttataat ggttacaaat	9600
aaagcaatag catcacaaat ttcacaaata aagcattttt ttcactgcat tctagttgtg	9660
gtttgtccaa actcatcaat gtatcttatc atgtctggct ctagctatcc cgcccctaac	9720
tccgcccata ccgcccctaa ctccgcccag ttccgcccata tctccgcccc atggctgact	9780
aatTTTTTTTT atttatgcag aggccgaggc cgccctcgcc tctgagctat tccagaagta	9840
gtgaggaggc ttttttggag gcctagggac gtacccaatt cgccctatag tgagtcgtat	9900
tacgcgcgct cactggccgt cgtttttacaa cgtcgtgact gggaaaaccc tggcgttacc	9960
caacttaatc gccttgagc acatccccct ttcgccagct ggcgtaatag cgaagaggcc	10020
cgcaccgatc gcccttccca acagttgcgc agcctgaatg gcgaatggga cgcgccctgt	10080
agcggcgcat taagcgcggc ggggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc	10140
agcgccctag cgcccgctcc tttcgctttc ttcccttcct ttctcgccac gttcgccggc	10200
tttccccgtc aagctctaaa tcggggggctc cctttagggt tccgatttag tgctttacgg	10260

cacctcgacc ccaaaaaaact tgattagggt gatggttcac gtagtgggcc atcgccctga	10320
tagacggttt ttcgcccttt gacgttggag tccacgttct ttaatagtgg actcttgttc	10380
caaactggaa caacactcaa ccctatctcg gtctattctt ttgatttata agggattttg	10440
ccgatttcgg cctattgggt aaaaaatgag ctgatttaac aaaaatttaa cgcgaaat	10500
aacaaaatat taacgcttac aatttaggtg gcacttttcg gggaaatgtg cgcggaaccc	10560
ctatttgttt atttttctaa atacattcaa atatgtatcc gctcatgaga caataaccct	10620
gataaatgct tcaataatag cacctagatc aagagacagg atgaggatcg tttcgcatga	10680
ttgaacaaga tggattgcac gcaggttctc cgcccgcttg ggtggagagg ctattcggct	10740
atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc	10800
aggggcgccc gggtcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcaag	10860
acgaggcagc gcggctatcg tggctggcca cgacgggcgt tccttgcgca gctgtgctcg	10920
acgttgtcac tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc	10980
tcctgtcatc tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc	11040
ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaac catcgcatcg	11100
agcgagcacg tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc	11160
atcaggggct cgcgccagcc gaactgttcg ccaggctcaa ggcgagcatg cccgacggcg	11220
aggatctcgt cgtgacccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc	11280
gcttttctgg attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag	11340
cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg	11400
tgctttacgg tatcgccgct cccgattcgc agc	11433

<210> 25
 <211> 1455
 <212> PRT
 <213> Homo sapiens
 <400> 25

Met	Ser	Asp	Ser	Trp	Val	Pro	Asn	Ser	Ala	Ser	Gly	Gln	Asp	Pro	Gly	1	5	10	15
Gly	Arg	Arg	Arg	Ala	Trp	Ala	Glu	Leu	Leu	Ala	Gly	Arg	Val	Lys	Arg	20	25	30	
Glu	Lys	Tyr	Asn	Pro	Glu	Arg	Ala	Gln	Lys	Leu	Lys	Glu	Ser	Ala	Val	35	40	45	
Arg	Leu	Leu	Arg	Ser	His	Gln	Asp	Leu	Asn	Ala	Leu	Leu	Leu	Glu	Val	50	55	60	
Glu	Gly	Pro	Leu	Cys	Lys	Lys	Leu	Ser	Leu	Ser	Lys	Val	Ile	Asp	Cys	65	70	75	80
Asp	Ser	Ser	Glu	Ala	Tyr	Ala	Asn	His	Ser	Ser	Ser	Phe	Ile	Gly	Ser	85	90	95	
Ala	Leu	Gln	Asp	Gln	Ala	Ser	Arg	Leu	Gly	Val	Pro	Val	Gly	Ile	Leu	100	105	110	
Ser	Ala	Gly	Met	Val	Ala	Ser	Ser	Val	Gly	Gln	Ile	Cys	Thr	Ala	Pro	115	120	125	
Ala	Glu	Thr	Ser	His	Pro	Val	Leu	Leu	Thr	Val	Glu	Gln	Arg	Lys	Lys	130	135	140	
Leu	Ser	Ser	Leu	Leu	Glu	Phe	Ala	Gln	Tyr	Leu	Leu	Ala	His	Ser	Met	145	150	155	160
Phe	Ser	Arg	Leu	Ser	Phe	Cys	Gln	Glu	Leu	Trp	Lys	Ile	Gln	Ser	Ser	165	170	175	
Leu	Leu	Leu	Glu	Ala	Val	Trp	His	Leu	His	Val	Gln	Gly	Ile	Val	Ser	180	185	190	

2017322511 10 Aug 2021

Leu	Gln	Glu	Leu	Leu	Glu	Ser	His	Pro	Asp	Met	His	Ala	Val	Gly	Ser
		195					200					205			
Trp	Leu	Phe	Arg	Asn	Leu	Cys	Cys	Leu	Cys	Glu	Gln	Met	Glu	Ala	Ser
	210					215					220				
Cys	Gln	His	Ala	Asp	Val	Ala	Arg	Ala	Met	Leu	Ser	Asp	Phe	Val	Gln
225					230					235					240
Met	Phe	Val	Leu	Arg	Gly	Phe	Gln	Lys	Asn	Ser	Asp	Leu	Arg	Arg	Thr
				245					250					255	
Val	Glu	Pro	Glu	Lys	Met	Pro	Gln	Val	Thr	Val	Asp	Val	Leu	Gln	Arg
			260					265					270		
Met	Leu	Ile	Phe	Ala	Leu	Asp	Ala	Leu	Ala	Ala	Gly	Val	Gln	Glu	Glu
		275					280					285			
Ser	Ser	Thr	His	Lys	Ile	Val	Arg	Cys	Trp	Phe	Gly	Val	Phe	Ser	Gly
	290					295					300				
His	Thr	Leu	Gly	Ser	Val	Ile	Ser	Thr	Asp	Pro	Leu	Lys	Arg	Phe	Phe
305					310					315					320
Ser	His	Thr	Leu	Thr	Gln	Ile	Leu	Thr	His	Ser	Pro	Val	Leu	Lys	Ala
				325					330					335	
Ser	Asp	Ala	Val	Gln	Met	Gln	Arg	Glu	Trp	Ser	Phe	Ala	Arg	Thr	His
			340					345					350		
Pro	Leu	Leu	Thr	Ser	Leu	Tyr	Arg	Arg	Leu	Phe	Val	Met	Leu	Ser	Ala
		355					360					365			
Glu	Glu	Leu	Val	Gly	His	Leu	Gln	Glu	Val	Leu	Glu	Thr	Gln	Glu	Val
	370					375					380				
His	Trp	Gln	Arg	Val	Leu	Ser	Phe	Val	Ser	Ala	Leu	Val	Val	Cys	Phe

10 Aug 2021
2017322511

385		390		395		400									
Pro	Glu	Ala	Gln	Gln	Leu	Leu	Glu	Asp	Trp	Val	Ala	Arg	Leu	Met	Ala
			405					410						415	
Gln	Ala	Phe	Glu	Ser	Cys	Gln	Leu	Asp	Ser	Met	Val	Thr	Ala	Phe	Leu
			420					425					430		
Val	Val	Arg	Gln	Ala	Ala	Leu	Glu	Gly	Pro	Ser	Ala	Phe	Leu	Ser	Tyr
		435					440					445			
Ala	Asp	Trp	Phe	Lys	Ala	Ser	Phe	Gly	Ser	Thr	Arg	Gly	Tyr	His	Gly
	450					455					460				
Cys	Ser	Lys	Lys	Ala	Leu	Val	Phe	Leu	Phe	Thr	Phe	Leu	Ser	Glu	Leu
465					470					475					480
Val	Pro	Phe	Glu	Ser	Pro	Arg	Tyr	Leu	Gln	Val	His	Ile	Leu	His	Pro
				485					490					495	
Pro	Leu	Val	Pro	Ser	Lys	Tyr	Arg	Ser	Leu	Leu	Thr	Asp	Tyr	Ile	Ser
			500					505					510		
Leu	Ala	Lys	Thr	Arg	Leu	Ala	Asp	Leu	Lys	Val	Ser	Ile	Glu	Asn	Met
		515					520					525			
Gly	Leu	Tyr	Glu	Asp	Leu	Ser	Ser	Ala	Gly	Asp	Ile	Thr	Glu	Pro	His
	530					535					540				
Ser	Gln	Ala	Leu	Gln	Asp	Val	Glu	Lys	Ala	Ile	Met	Val	Phe	Glu	His
545					550					555					560
Thr	Gly	Asn	Ile	Pro	Val	Thr	Val	Met	Glu	Ala	Ser	Ile	Phe	Arg	Arg
				565					570					575	
Pro	Tyr	Tyr	Val	Ser	His	Phe	Leu	Pro	Ala	Leu	Leu	Thr	Pro	Arg	Val
			580					585					590		

10 Aug 2021
2017322511

Leu Pro Lys Val Pro Asp Ser Arg Val Ala Phe Ile Glu Ser Leu Lys
595 600 605

Arg Ala Asp Lys Ile Pro Pro Ser Leu Tyr Ser Thr Tyr Cys Gln Ala
610 615 620

Cys Ser Ala Ala Glu Glu Lys Pro Glu Asp Ala Ala Leu Gly Val Arg
625 630 635 640

Ala Glu Pro Asn Ser Ala Glu Glu Pro Leu Gly Gln Leu Thr Ala Ala
645 650 655

Leu Gly Glu Leu Arg Ala Ser Met Thr Asp Pro Ser Gln Arg Asp Val
660 665 670

Ile Ser Ala Gln Val Ala Val Ile Ser Glu Arg Leu Arg Ala Val Leu
675 680 685

Gly His Asn Glu Asp Asp Ser Ser Val Glu Ile Ser Lys Ile Gln Leu
690 695 700

Ser Ile Asn Thr Pro Arg Leu Glu Pro Arg Glu His Ile Ala Val Asp
705 710 715 720

Leu Leu Leu Thr Ser Phe Cys Gln Asn Leu Met Ala Ala Ser Ser Val
725 730 735

Ala Pro Pro Glu Arg Gln Gly Pro Trp Ala Ala Leu Phe Val Arg Thr
740 745 750

Met Cys Gly Arg Val Leu Pro Ala Val Leu Thr Arg Leu Cys Gln Leu
755 760 765

Leu Arg His Gln Gly Pro Ser Leu Ser Ala Pro His Val Leu Gly Leu
770 775 780

Ala	Ala	Leu	Ala	Val	His	Leu	Gly	Glu	Ser	Arg	Ser	Ala	Leu	Pro	Glu
785					790					795					800
Val	Asp	Val	Gly	Pro	Pro	Ala	Pro	Gly	Ala	Gly	Leu	Pro	Val	Pro	Ala
				805					810					815	
Leu	Phe	Asp	Ser	Leu	Leu	Thr	Cys	Arg	Thr	Arg	Asp	Ser	Leu	Phe	Phe
			820					825					830		
Cys	Leu	Lys	Phe	Cys	Thr	Ala	Ala	Ile	Ser	Tyr	Ser	Leu	Cys	Lys	Phe
		835					840					845			
Ser	Ser	Gln	Ser	Arg	Asp	Thr	Leu	Cys	Ser	Cys	Leu	Ser	Pro	Gly	Leu
	850					855					860				
Ile	Lys	Lys	Phe	Gln	Phe	Leu	Met	Phe	Arg	Leu	Phe	Ser	Glu	Ala	Arg
865					870					875					880
Gln	Pro	Leu	Ser	Glu	Glu	Asp	Val	Ala	Ser	Leu	Ser	Trp	Arg	Pro	Leu
				885					890					895	
His	Leu	Pro	Ser	Ala	Asp	Trp	Gln	Arg	Ala	Ala	Leu	Ser	Leu	Trp	Thr
			900					905					910		
His	Arg	Thr	Phe	Arg	Glu	Val	Leu	Lys	Glu	Glu	Asp	Val	His	Leu	Thr
		915					920					925			
Tyr	Gln	Asp	Trp	Leu	His	Leu	Glu	Leu	Glu	Ile	Gln	Pro	Glu	Ala	Asp
	930					935					940				
Ala	Leu	Ser	Asp	Thr	Glu	Arg	Gln	Asp	Phe	His	Gln	Trp	Ala	Ile	His
945					950					955					960
Glu	His	Phe	Leu	Pro	Glu	Ser	Ser	Ala	Ser	Gly	Gly	Cys	Asp	Gly	Asp
				965					970					975	

2017322511 10 Aug 2021

Leu	Gln	Ala	Ala	Cys	Thr	Ile	Leu	Val	Asn	Ala	Leu	Met	Asp	Phe	His	
			980					985					990			
Gln	Ser	Ser	Arg	Ser	Tyr	Asp	His	Ser	Glu	Asn	Ser	Asp	Leu	Val	Phe	
		995					1000					1005				
Gly	Gly	Arg	Thr	Gly	Asn	Glu	Asp	Ile	Ile	Ser	Arg	Leu	Gln	Glu		
	1010					1015					1020					
Met	Val	Ala	Asp	Leu	Glu	Leu	Gln	Gln	Asp	Leu	Ile	Val	Pro	Leu		
	1025					1030					1035					
Gly	His	Thr	Pro	Ser	Gln	Glu	His	Phe	Leu	Phe	Glu	Ile	Phe	Arg		
	1040					1045					1050					
Arg	Arg	Leu	Gln	Ala	Leu	Thr	Ser	Gly	Trp	Ser	Val	Ala	Ala	Ser		
	1055					1060					1065					
Leu	Gln	Arg	Gln	Arg	Glu	Leu	Leu	Met	Tyr	Lys	Arg	Ile	Leu	Leu		
	1070					1075					1080					
Arg	Leu	Pro	Ser	Ser	Val	Leu	Cys	Gly	Ser	Ser	Phe	Gln	Ala	Glu		
	1085					1090					1095					
Gln	Pro	Ile	Thr	Ala	Arg	Cys	Glu	Gln	Phe	Phe	His	Leu	Val	Asn		
	1100					1105					1110					
Ser	Glu	Met	Arg	Asn	Phe	Cys	Ser	His	Gly	Gly	Ala	Leu	Thr	Gln		
	1115					1120					1125					
Asp	Ile	Thr	Ala	His	Phe	Phe	Arg	Gly	Leu	Leu	Asn	Ala	Cys	Leu		
	1130					1135					1140					
Arg	Ser	Arg	Asp	Pro	Ser	Leu	Met	Val	Asp	Phe	Ile	Leu	Ala	Lys		
	1145					1150					1155					
Cys	Gln	Thr	Lys	Cys	Pro	Leu	Ile	Leu	Thr	Ser	Ala	Leu	Val	Trp		

10 Aug 2021
2017322511

1160		1165		1170
Trp Pro 1175	Ser Leu Glu Pro	Val 1180	Leu Leu Cys Arg	Trp 1185 Arg Arg His
Cys Gln 1190	Ser Pro Leu Pro	Arg 1195	Glu Leu Gln Lys	Leu 1200 Gln Glu Gly
Arg Gln 1205	Phe Ala Ser Asp	Phe 1210	Leu Ser Pro Glu	Ala 1215 Ala Ser Pro
Ala Pro 1220	Asn Pro Asp Trp	Leu 1225	Ser Ala Ala Ala	Leu 1230 His Phe Ala
Ile Gln 1235	Gln Val Arg Glu	Glu 1240	Asn Ile Arg Lys	Gln 1245 Leu Lys Lys
Leu Asp 1250	Cys Glu Arg Glu	Glu 1255	Leu Leu Val Phe	Leu 1260 Phe Phe Phe
Ser Leu 1265	Met Gly Leu Leu	Ser 1270	Ser His Leu Thr	Ser 1275 Asn Ser Thr
Thr Asp 1280	Leu Pro Lys Ala	Phe 1285	His Val Cys Ala	Ala 1290 Ile Leu Glu
Cys Leu 1295	Glu Lys Arg Lys	Ile 1300	Ser Trp Leu Ala	Leu 1305 Phe Gln Leu
Thr Glu 1310	Ser Asp Leu Arg	Leu 1315	Gly Arg Leu Leu	Leu 1320 Arg Val Ala
Pro Asp 1325	Gln His Thr Arg	Leu 1330	Leu Pro Phe Ala	Phe 1335 Tyr Ser Leu
Leu Ser 1340	Tyr Phe His Glu	Asp 1345	Ala Ala Ile Arg	Glu 1350 Glu Ala Phe

10 Aug 2021
2017322511

Leu 1355	His	Val	Ala	Val	Asp	Met 1360	Tyr	Leu	Lys	Leu	Val 1365	Gln	Leu	Phe
Val 1370	Ala	Gly	Asp	Thr	Ser	Thr 1375	Val	Ser	Pro	Pro	Ala 1380	Gly	Arg	Ser
Leu 1385	Glu	Leu	Lys	Gly	Gln	Gly 1390	Asn	Pro	Val	Glu	Leu 1395	Ile	Thr	Lys
Ala 1400	Arg	Leu	Phe	Leu	Leu	Gln 1405	Leu	Ile	Pro	Arg	Cys 1410	Pro	Lys	Lys
Ser 1415	Phe	Ser	His	Val	Ala	Glu 1420	Leu	Leu	Ala	Asp	Arg 1425	Gly	Asp	Cys
Asp 1430	Pro	Glu	Val	Ser	Ala	Ala 1435	Leu	Gln	Ser	Arg	Gln 1440	Gln	Ala	Ala
Pro 1445	Asp	Ala	Asp	Leu	Ser	Gln 1450	Glu	Pro	His	Leu	Phe 1455			